

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:00:20 ; Search time 13.566 Seconds  
(without alignments)  
339.796 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586

Perfect score: 332

Sequence: 1 CSLSNAGHCSGDPCCNNNTSC.....CDITEYCTGD;GQCPEPLMK 55

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283306

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	213	64.2	524	2	S38539	disintegrin-like m
2	213	64.2	524	2	S38539	disintegrin-like m
3	168.5	50.8	814	2	G02330	disintegrin-like m
4	168	50.6	617	2	S48160	metalloproteinase
5	153	46.1	357	2	S23403	sperm surface prot
6	153	46.1	736	2	S47645	WDC I protein - c
7	152	45.8	734	2	JC4861	ferillin beta cha
8	151	45.5	735	2	G02937	ferillin beta - cr
9	148	44.6	600	2	I49281	ferillin alpha pre
10	147	44.3	616	2	A55796	ecarin precursor -
11	146.5	44.1	419	2	A59474	metalloproteinase
12	145.5	43.8	571	2	S24789	parahagrin C precu
13	145.5	43.8	609	2	S55270	catrocollastatin p
14	144	43.4	825	2	S55060	ferillin alpha-IT
15	144	43.4	905	2	S55059	ferillin alpha-I-
16	143	43.1	150	2	S60258	meltrin beta - mou
17	136	41.0	473	2	I49283	ADAM 4 protein pre
18	136	41.0	549	2	S48169	metalloproteinase
19	136	41.0	823	2	S18968	vascular apoptosis
20	135	40.7	610	2	JC7530	hemorrhagic protei
21	134	40.4	416	2	A57877	meltrin alpha - mo
22	134	40.4	903	2	S60257	disintegrin and me
23	134	40.1	952	2	T18900	platelet aggregated
24	133	39.8	216	2	JX0265	androgen-regulated
25	132	39.2	789	2	S28259	disintegrin and me
26	130	39.2	655	2	JC7850	metalloproteinase
27	130	39.2	660	2	ST1949	metalloproteinase
28	130	39.2	732	2	I52361	platelet aggregated
29	129.5	39.0	209	2	JX0266	platelet aggregated

30	129	38.9	735	2	I48101	ADAM 6 protein pre
31	128	38.6	777	2	I48100	ADAM 5 protein pre
32	127.5	38.4	719	2	S41607	acetylserin A (EC 3.
33	122	36.7	429	2	A42972	coagulation factor
34	122	36.7	713	2	I65253	disintegrin-like t
35	121.5	36.6	73	2	D35982	trigramin beta-2 -
36	121.5	36.6	478	2	JC4880	fibronolytic metal
37	121	36.4	776	2	S28258	androgen-regulated
38	120.5	36.3	1042	2	T26644	hypothetical prote
39	119	35.8	826	2	A60385	monocyte surface a
40	117	35.2	151	2	S60259	meltrin gamma - mo
41	115	34.6	756	2	S47656	WDC II protein -
42	113.5	34.2	71	2	A59411	platelet aggregated
43	113.5	34.2	73	2	A59410	platelet aggregated
44	113	34.0	512	2	T37819	probable zinc meta
45	112.5	33.9	71	2	A59409	platelet aggregated

#### ALIGNMENTS

##### RESULT 1

S38539 disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 1 - human

N:Alternate names: ADAM11; MDC

C/Species: Homo sapiens (man)

C/Date: 07-Oct-1994 #sequence\_revision 17-Nov-1995 #text\_change 26-May-2000

C/Accession: S38539; 152965 #sequence\_revision 17-Nov-1995 #text\_change 26-May-2000

R/Eml, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, J.; Ito, I.; Kasumi, F.; Nakamura, Y.

A/Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearrang

A/Reference number: S38539; WUID:94073190; PMID:8252040

A/Accession: S38539

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-524 <EMI>

A/Cross-references: GB:D17390; NID:9452188; PIDN:BA04213.1; PID:di004732; PID:9484255

R/Katagiri, T.; Harada, Y.; Eml, M.; Nakamura, Y.

C/Genet. Cell Genet. 68, 39-44, 1995

A/Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and

A/Reference number: 152965; WUID:95044425; PMID:7956356

A/Accession: 152965

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Residues: 7-225 'D', 227-524 <KAT>

A/Cross-references: GB:D181872; NID:9505090; PIDN:BA06671.1; PID:di007243; PID:9836684

C/Comment: For an alternative splice form, see PIR:165967.

C/Genetics:

A/Gene: GDB:MDC; ADAM11

A/Cross-references: GDB:230267; OMIM:155120

A/Map position: 17q21.3-17q21.3

A/Intons: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341/

C/Superfamily: disintegrin homology

C/Keywords: alternative splicing; hydrolase; metalloproteinase

F:344-427/Domain: disintegrin homology <DIS>

Query Match 64.2%; Score 213; DB 2; Length 524;

Best Local Similarity 65.5%; Pred. No. 4.9e-14;

Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Qy 1 CSLSNAGHCSGDPCCNNNTSCLEFQPRGYECRDVNECDITEYCTGDSGQCPEPLMK 55  
Db 378 CTTTHDMCSGDCRCR--CKYEPREVSCREAVNECDIATCTGDSGQCPEPLMK 430

##### RESULT 2

I65967 disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 2 - human

N:Alternate names: ADAM11; MDC

C/Species: Homo sapiens (man)

C/Date: 29-May-1998 #sequence\_revision 17-Mar-2000 #text\_change 11-Jan-2002

C/Accession: I65967; S38539

R/Katagiri, T.; Harada, Y.; Eml, M.; Nakamura, Y.

Cytogenet. Cell Genet. 68, 39-44, 1995  
A:Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and  
A:Reference number: 152965; MUID:95044425; PMID:7956356  
A:Accession: 165967  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 7-670 <KAT>  
A:Cross-references: GB:D31872; NID:G505090; PIDN:BA06670.1; PID:G836683  
R:Emi, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, J.; Ito, I.; Kasumi, F.; Nakamura  
Nature Genet. 5, 151-157, 1993  
A:Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearran  
A:Reference number: S88539; MUID:94073190; PMID:8252040  
A:Accession: S88539  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-225, 'N', 227-495, 'POGRAVWLPRLCOHLMSSARPGGRHO' <EMI>  
A:Cross-references: GB:D17390; NID:G452188; PIDN:BA04213.1; PID:G484255  
C:Comment: For an alternative splice form, see PIR:S88539.  
C:Genetics:  
A:Gene: GDB:MDC, ADAM11  
A:Cross-references: GDB:230267; OMIM:155120  
A:Map position: 17q21.3-17q21.3  
A:Introns: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341  
C:Superfamily: disintegrin homology  
C:Keywords: alternative splicing; hydrolase; metalloproteinase  
F:344-427/Domain: disintegrin homology <DIS>

Query Match 64.2%; Score 213; DB 2; Length 670;  
Best Local Similarity 65.5%; Pred. No. 5, 9e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Qy 1 CSLSNGAHCSGDPCCNNTSCLPQPRGYECRDVAVECDITEXCTGDSQCPCPNLHK 55  
Db 378 CTLTHDAMCSGDLCCRR--CKYPRGVSCRAVNECDIAETCTGDSQCPCPNLHK 430

RESULT 3  
G02390  
disintegrin-like metalloproteinase MDC15 (EC 3.4.24.-) - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 31-Dec-2000  
C:Accession: G02390; PC4263  
R:Herren, B.; Raines, E.W.; Ross, R.  
submitted to the EMBL Data Library, January 1996  
A:Reference number: H01157  
A:Accession: G02390  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-814 <HER>  
A:Cross-references: EMBL:U46005; NID:G1335871; PIDN:ACG51112.1; PID:G1335872  
R:McKie, N.; Edwards, T.; Dallas, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Russell, B.  
Biochem. Biophys. Res. Commun. 230, 335-339, 1997  
A:Title: Expression of members of a novel membrane linked metalloproteinase family (ADAM  
A:Reference number: PC4263; MUID:9716871; PMID:9016778  
A:Accession: PC4263  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-461 <MCK>  
A:Experimental source: articular chondrocyte  
C:Comment: This protein is a membrane bound protein and involved in cell/cell and cell/m  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: hydrolase; metalloproteinase; zinc  
F:420-503/Domain: disintegrin homology <DIS>  
F:348,352,358/Binding site: zinc (His) #status predicted  
F:349/Active site: Glu #status predicted

Query Match 50.8%; Score 168.5; DB 2; Length 814;  
Best Local Similarity 51.9%; Pred. No. 1, 8e-09;  
Matches 28; Conservative 8; Mismatches 15; Indels 3; Gaps 2;

Qy 1 CSLSNGAHCSGDPCCNNTSCLPQPRGYECRDVAVECDITEXCTGDSQCPCPNL 53  
Db 453 CQLRPGACASDGPCCN--CQLRPSGWGCRPTRGDCDLPFCQDSQCCPDPV 504

RESULT 4  
S48160  
metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper  
N:Contains: disintegrin  
C:Species: Echis pyramicum leakeyi  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 21-Jul-2000  
A:Accession: S48160  
R:Paine, M.J.I.; Moura-da-Silva, A.M.; Theakston, R.D.G.; Crampson, J.M.  
Eur. J. Biochem. 224, 483-488, 1994  
A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramicum leakeyi).  
A:Reference number: S48160; MUID:95010025; PMID:7925363  
A:Accession: S48160  
A:Molecule type: mRNA  
A:Residues: 1-617 <PAI>  
A:Cross-references: GB:Y78970; NID:G763092; PIDN:CAA55565.1; PID:G763093  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: hydrolase; metalloproteinase; venom  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-192/Domain: propeptide #status predicted <PRO>  
F:193-617/Domain: metalloproteinase H-I #status predicted <MAT>  
F:403-485/Domain: disintegrin homology <DIS>  
F:337/Active site: Glu #status predicted

Query Match 50.6%; Score 168; DB 2; Length 617;  
Best Local Similarity 52.0%; Pred. No. 1, 7e-09;  
Matches 26; Conservative 9; Mismatches 13; Indels 2; Gaps 1;

Qy 1 CSLSNGAHCSGDPCCNNTSCLPQPRGYECRDVAVECDITEXCTGDSQCPCPNLH 50  
Db 436 CKLTPSGQCADGECN--OCFRFAPRTCCRRKIDCDVPEYCTGDSGCECP 483

RESULT 5  
S23403  
sperm surface protein PH-30 beta chain precursor - guinea pig (fragment)  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 09-Jun-2000  
C:Accession: S23403; S25696  
R:Boebel, C.P.; Wolfberg, T.G.; Turck, C.W.; Myles, D.G.; Primakoff, P.; White, J.M.  
Nature 356, 248-252, 1992  
A:Title: A potential fusion peptide and an integrin ligand domain in a protein active in  
A:Reference number: S23402; MUID:92204234; PMID:1552944  
A:Accession: S23403  
A:Molecule type: mRNA  
A:Residues: 1-357 <BOI>  
A:Cross-references: EMBL:Z11720  
A:Accession: S25696  
A:Molecule type: protein  
A:Residues: 5-8, 'X', 10-32:67-88;125-129, 'X', 131-134, 'X', 136-141, 'X', 143;154, 'X', 156-161, 'J'  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-/Domain: signal sequence (fragment) #status predicted <SIG>  
F:5-357/Product: sperm surface protein PH-30 beta chain #status experimental <MAT>  
F:5-87/Domain: disintegrin homology <DIS>  
F:303-323/Domain: transmembrane #status predicted <TM>  
F:78,186,256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.1%; Score 153; DB 2; Length 357;  
Best Local Similarity 44.4%; Pred. No. 3, 4e-08;  
Matches 24; Conservative 11; Mismatches 17; Indels 2; Gaps 1;

Qy 1 CSLSNGAHCSGDPCCNNTSCLPQPRGYECRDVAVECDITEXCTGDSQCPCPNLH 54  
Db 38 CRLKTSRCAQGPCN--QCEFKTGEVCRESTDECDLPYCNWSSGACQCEDLV 89

RESULT 6  
S47645  
TMDC I protein - crab-eating macaque  
C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 21-Jul-2000

C:Accession: S47645  
 R:Barker, H.L.; Perry, A.C.F.; Jones, R.; Hall, L.  
 C:Species: Homo sapiens (man)  
 A:Title: Sequence and expression of a monkey testicular transcript encoding tMDC 1, a novel protein.  
 A:Reference number: S47645; MUID:9432353; PMID:8049267  
 A:Accession: S47645  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-736 <BAR>  
 A:Cross-references: EMBL:X76637; NID:G535016; PIDN:CA54083.1; PID:G535017  
 C:Superfamily: mouse meltrin alpha; disintegrin homology  
 F:392-477/Domain: disintegrin homology <DIS>

Query Match 46.1%; Score 153; DB 2; Length 736;  
 Best Local Similarity 45.3%; Pred. No. 5.9e-08;  
 Matches 24; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Db 1 CSLSNGAHGSDGPCNNNTSCLFOPRGYECRDVAVECDITCYCTGDSGCCPPN 53  
 426 CTLIGFAECGSGPCNNKCTIYAGHVCRRKSIDCDPEYCNCTSEFCVDPV 478

RESULT 7  
 fertilin beta chain - human  
 C:Species: Homo sapiens (man)  
 C>Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 21-Jan-2000  
 A:Accession: J04861  
 R:Gupta, S.K.; Alves, K.; O'Neil Palladino, L.; Mark, G.E.; Hollis, G.F.  
 Biochem. Biophys. Res. Commun. 224, 318-326, 1996  
 A:Title: Molecular cloning of the human fertilin beta subunit.  
 A:Reference number: J04861; MUID:96295488; PMID:8702389  
 A:Accession: J04861  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-734 <GUP>  
 A:Cross-references: GB:U38805; NID:G4151118; PIDN:AAD04206.1; PID:G4151119  
 C:Comment: This protein is an integral sperm membrane glycoprotein, and plays a role in fertilization.  
 C:Superfamily: mouse meltrin alpha; disintegrin homology  
 C:Keywords: glycoprotein; integrin binding; transmembrane protein  
 F:382-734/Product: fertilin beta chain #status predicted <NAT>  
 F:382-467/Domain: disintegrin homology <DIS>  
 F:448-450/Region: integrin binding #status predicted  
 F:688-708/Domain: transmembrane #status predicted <TM>  
 F:121,219,352,458,565/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.8%; Score 152; DB 2; Length 734;  
 Best Local Similarity 46.2%; Pred. No. 7.5e-08;  
 Matches 24; Conservative 7; Mismatches 19; Indels 2; Gaps 1;

Db 1 CSLSNGAHGSDGPCNNNTSCLFOPRGYECRDVAVECDITCYCTGDSGCCPPN 52  
 418 CRFAGSGNCAEGPCEN--CLFMSKERNCRPSFECDLPYCNSSASCPEN 467

RESULT 8  
 fertilin beta - crab-eating macaque  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 01-Dec-2000  
 A:Accession: G02937; S55061  
 R:Ramirez, C.S.; Myles, D.G.; White, J.M.; Primakoff, P.  
 submitted to the EMBL Data Library, August 1995  
 A:Reference number: G12615  
 A:Accession: G02937  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-735 <RAM>  
 A:Cross-references: EMBL:U33959; NID:G998339; PID:G998340  
 Biochem. J. 307, 843-850, 1995  
 A:Title: Cloning and analysis of monkey fertilin reveals novel alpha subunit isoforms.  
 A:Reference number: S55059; MUID:95260313; PMID:7741716  
 A:Accession: S55061

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-722, 'S', 724-735 <PER>  
 A:Cross-references: EMBL:X77653; NID:G794076; PIDN:CA54733.1; PID:G794077  
 C:Superfamily: mouse meltrin alpha; disintegrin homology  
 F:383-468/Domain: disintegrin homology <DIS>

Query Match 45.5%; Score 151; DB 2; Length 735;  
 Best Local Similarity 46.2%; Pred. No. 9.4e-08;  
 Matches 24; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

Db 1 CSLSNGAHGSDGPCNNNTSCLFOPRGYECRDVAVECDITCYCTGDSGCCPPN 52  
 419 CRFAGSGNCAEGPCEN--CLFMSQERVCRRPSFECDLPYCNCTSEFCVDPV 468

RESULT 9  
 fertilin alpha precursor - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jun-2000  
 A:Accession: J49281  
 R:Wolfsberg, T.G.; Straight, P.D.; Gerena, R.L.; Huovila, A.P.; Primakoff, P.; Myles, D.  
 Dev. Biol. 169, 378-383, 1995  
 A:Title: ADAM, a widely distributed and developmentally regulated gene family encoding m  
 A:Reference number: J49281; MUID:95269891; PMID:7750654  
 A:Accession: J49281  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-600 <RNS>  
 A:Cross-references: EMBL:U22056; NID:G965009; PIDN:AAA74920.1; PID:G965010  
 C:Genetics:  
 C:Superfamily: unassigned disintegrins; disintegrin homology  
 F:246-356/Domain: disintegrin homology <DIS>  
 F:180/Active site: Gu #status predicted

Query Match 44.6%; Score 148; DB 2; Length 600;  
 Best Local Similarity 48.1%; Pred. No. 1.6e-07;  
 Matches 25; Conservative 8; Mismatches 17; Indels 2; Gaps 1;

Db 1 CSLSNGAHGSDGPCNNNTSCLFOPRGYECRDVAVECDITCYCTGDSGCCPPN 52  
 277 CTKLEGAQCSGLCCV--CTFKKGSICRPABDVCDLPYCNCTSEFCVDPV 326

RESULT 10  
 ecarin precursor - saw-scaled viper  
 C:Species: Echis carinatus (saw-scaled viper)  
 C>Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 21-Jul-2000  
 A:Accession: A55796  
 R:Nishida, S.; Fujita, T.; Kohno, N.; Atoda, H.; Morita, T.; Takeya, H.; Kido, I.; Paine  
 Biochemistry 34, 1771-1778, 1995  
 A:Title: cDNA cloning and deduced amino acid sequence of prothrombin activator (ecarin)  
 A:Reference number: A55796; MUID:95151760; PMID:7849037  
 A:Accession: A55796  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-616 <NIS>  
 A:Cross-references: GB:J32212; NID:G717090; PIDN:BA06910.1; PID:G717091  
 C:Superfamily: mouse meltrin alpha; disintegrin homology  
 F:404-486/Domain: disintegrin homology <DIS>  
 F:338/Active site: Glu #status predicted

Query Match 44.3%; Score 147; DB 2; Length 616;  
 Best Local Similarity 46.2%; Pred. No. 2.1e-07;  
 Matches 24; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

Db 1 CSLSNGAHGSDGPCNNNTSCLFOPRGYECRDVAVECDITCYCTGDSGCCPPN 52  
 437 CKLRGAGCGGEC--CKIRKAGTECRPARDDCVABCTGQASACPEN 486

Mon Oct 27 15:28:08 2003

us-09-634-252a-4\_copy\_532\_586.rpr

Page 4

Search completed: October 21, 2003, 10:11:34  
Job time : 16.566 secs

---





FT Region 313..540  
 XX /note="Fc region"  
 XX  
 PN WO200162905-A2.  
 PD 30-AUG-2001.  
 PF 23-FEB-2001; 2001WO-US05701.  
 XX  
 PR 25-FEB-2000; 2000US-184865P.  
 XX  
 PA (IMMUNEX ) IMMUNEX CORP.  
 XX  
 PI Farnsworth WC, Cerretti DP, Poindexter KM, Black RA;  
 XX  
 DR WPI: 2001-625725/72.  
 XX  
 DR N-PSDB; AAD21439.  
 XX  
 PT Antagonizing the binding of an integrin to its ligand useful for the  
 PT treatment of angiogenesis comprises administration of an  
 PT ADAM-disintegrin domain polypeptide  
 XX  
 PS Claim 11; Page 60-61; 66pp; English.  
 XX  
 CC The invention relates to the method and use of ADAM disintegrin domain  
 CC polypeptides for inhibiting the biological activity of integrins,  
 CC endothelial cell migration and angiogenesis. ADAM disintegrin domain  
 CC polypeptides are used for treatment of ocular disorders, malignant and  
 CC metastatic conditions, inflammatory diseases, osteoporosis and other  
 CC conditions mediated by accelerated bone resorption, stenosis,  
 CC inappropriate platelet activation, recruitment or aggregation, thrombosis  
 CC or a condition requiring aggregation, thrombosis or a condition requiring  
 CC tissue repair or wound healing, angiogenesis, ocular neovascularization  
 CC or solid tumor, diabetic retinopathy, retinopathy or prematurity,  
 CC neovascular glaucoma, retinoblastoma, retrolental fibroplasias, rubeosis,  
 CC uveitis, macular degeneration and corneal graft neovascularization,  
 CC inflammatory diseases, ocular tumors, diseases associated with chorioidal  
 CC or iris neovascularization, arthritis, rheumatism, inflammatory bowel  
 CC disease, psoriasis, coronary artery disease or injury, myocardial  
 CC infarction or injury following myocardial infarction, stroke, unstable  
 CC angina, atherosclerosis, arteriosclerosis, pre-eclampsia, embolism,  
 CC platelet-associated ischaemic disorders including lung ischaemia,  
 CC coronary ischaemia, cerebral ischaemia, restenosis following percutaneous  
 CC coronary intervention including angioplasty, atherectomy, stent placement  
 CC and bypass surgery, thrombotic disorders including coronary artery  
 CC thrombosis, cerebral artery thrombosis, intracardiac thrombosis,  
 CC peripheral artery thrombosis, venous thrombosis, thrombosis and  
 CC coagulopathies associated with exposure to a foreign or injured tissue  
 CC surface and reocclusion following thrombosis, deep venous thrombosis,  
 CC pulmonary embolism, transient ischaemic attacks and another conditions  
 CC where vascular occlusion is a common underlying feature, in individuals  
 CC at high risk for thrombus formation of reformation, advanced coronary  
 CC artery disease, or for occlusion, reocclusion, stenosis and/or restenosis  
 CC of blood vessels or stroke benign tumors and preneoplastic conditions,  
 CC myocardial angiogenesis, haemophilic joints, scleroderma, vascular  
 CC adhesions, asthma and allergy, eczema and dermatitis, graft versus host  
 CC disease, sepsis, adult respiratory distress syndrome, telangiectasia and  
 CC wound granulation. The method are used in combination with angioplasty  
 CC procedures, such as balloon angioplasty, laser angioplasty, coronary  
 CC atherectomy or similar techniques, carotid endarterectomy, anastomosis of  
 CC vascular grafts, surgery having a high risk of thrombus formation (i.e.  
 CC coronary bypass surgery, insertion of a prosthetic valve or vessel and  
 CC the like), atherectomy, stent placement, placement of a chronic  
 CC cardiovascular device such as an in-dwelling catheter or prosthetic valve  
 CC or vessel, organ transplantation or bypass surgery. The present sequence  
 CC is ADAM disintegrin domain polypeptide fusion construct. The fusion  
 CC construct comprises of immunoglobulin K (IgK) leader, human ADAM  
 CC disintegrin and Fc region.  
 XX  
 SQ Sequence 540 AA;  
 Query Match 100.0%; Score 332; DB 22; Length 540;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-25;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSLSGAHCSGPPCCNNNTSCLPFGRGYECRPAVNECDITECTGDSGCCPNNLAK 55  
 DB 50 CSLSGAHCSGPPCCNNNTSCLPFGRGYECRPAVNECDITECTGDSGCCPNNLAK 104  
 RESULT 2  
 AAY54457  
 ID AAY54457 standard; Protein: 696 AA.  
 XX  
 AC AAY54457;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Amino acid sequence of a disintegrin homologue designated zdint1.  
 KW Human; disintegrin homologue; zdint1; cardiac myocyte; adipocyte;  
 KW gene therapy; cell-cell interaction; chondro sarcoma; atherosclerosis;  
 KW Alzheimer's disease; restenosis; ischemic reperfusion; obesity;  
 KW intimal hyperplasia; tumor; platelet aggregation; apoptosis;  
 KW neurogenesis; myogenesis; arthritis; myogenesis; neurogenesis;  
 KW connective tissue disorder; chondrogenesis; tumor proliferation;  
 KW inflammation.  
 KM  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH Domain 1..163  
 FT /note="propeptide sequence"  
 FT 164..382  
 FT Domain /note="protease sequence; this sequence is  
 FT specifically claimed in claim 4"  
 FT 383..464  
 FT Domain /note="disintegrin sequence; this sequence is  
 FT specifically claimed in claim 5"  
 FT Misc-difference 437..450  
 FT /note="this sequence is specifically claimed in claim 2"  
 FT 438..449  
 FT Region /note="disintegrin loop"  
 FT 443..445  
 FT Region /note="RGD binding loop"  
 FT 465..696  
 FT Domain /note="cysteine-rich domain; this sequence is  
 FT specifically claimed in claim 6"  
 FT  
 PN WO200002912-A2.  
 XX  
 PD 20-JAN-2000.  
 XX  
 PF 09-JUL-1999; 99WO-US15638.  
 XX  
 PR 10-JUL-1998; 98US-0113883.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Sheppard PO, Baindur N, Deisher TA, Bishop PD;  
 DR WPI: 2000-160898/14.  
 DR N-PSDB; AAZ45826.  
 XX  
 PT Polypeptide useful in modulating cell-cell interaction in tissues of  
 PT heart, brain, spinal cord and treating chondro sarcoma,  
 PT atherosclerosis, restenosis, obesity, intimal hyperplasia and tumors -  
 XX  
 PS Claim 7; Page 127-129; 132pp; English.  
 XX  
 CC The present sequence represents a human disintegrin homologue, designated  
 CC zdint1. The zdint1 polypeptide is a cardiac myocyte proliferation and  
 CC differentiation stimulator, as well as an adipocyte proliferation and  
 CC differentiation inhibitor. Polynucleotides encoding zdint1 are used in  
 CC gene therapy. The zdint1 polypeptide is useful in modulating cell-cell  
 CC interactions of cells derived from tissues of heart, brain, spinal cord

CC and skeletal muscle. It is useful in treating and diagnosing chondro  
CC sarcoma, atherosclerosis, Alzheimer's disease, restenosis, ischemic  
CC reperfusion, obesity, intimal hyperplasia and tumors of heart, brain,  
CC and spinal cord. The zdin1 polypeptide is also useful in identifying  
CC its new family members, antagonists, agonists and antibodies.  
CC Antagonists, antibodies and fusion proteins of zdin1 are useful in  
CC inhibiting platelet aggregation, apoptosis, neurogenesis and myogenesis.  
CC Agonists and antagonists are useful in studying cell-cell interactions,  
CC arthritis, myogenesis, neurogenesis, connective tissue disorders,  
CC chondrogenesis, tumour proliferation and suppression, extracellular  
CC matrix proteins, repair and remodelling of ischemic reperfusion,  
CC inflammation, and apoptosis.  
XX  
SQ Sequence 696 AA;  
Query Match 100.0%; Score 332; DB 21; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVECDITETCTGSGCCPPNLHK 55  
Db 409 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVECDITETCTGSGCCPPNLHK 463  
RESULT 3  
AAU07190 ID AAU07190 standard; Protein; 696 AA.  
XX  
XX AAU07190;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human disintegrin protease, zdin1.  
XX  
KM Human, disintegrin: zdin1; antiangiogenic; vascular; thrombolytic.  
KM cell matrix; cell-cell interactions; apoptosis; neurogenesis;  
KM connective tissue disorders; chondrogenesis; arthritis;  
KM tumour proliferation; ischaemia reperfusion; inflammation;  
KM chromosome 2q33.  
XX  
XX OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1..163  
FT /label= Propeptide  
FT 164..696  
FT Protein /label= Mature\_zdin1  
FT /note= "This sequence is specifically claimed in  
FT claim 4"  
FT Region /label= Protease region  
FT 164..382  
FT /label= "This sequence is specifically claimed in  
FT claim 1"  
FT Region /label= Disintegrin sequence  
FT 383..464  
FT /note= "Disintegrin sequence  
FT /note= "This sequence is specifically claimed in  
FT claim 2"  
FT 383..696  
FT /label= Disintegrin/cysteine-rich sequence  
FT /note= "This sequence is specifically claimed in  
FT claim 3"  
FT Region 438..449  
FT /label= Disintegrin\_loop  
FT 443..445  
FT Binding-site /label= RGD\_binding\_loop  
XX  
XX US6265199-B1.  
XX  
XX 24-JUL-2001.  
XX  
XX 09-JUL-1999; 99US-0351414.  
XX  
XX 10-JUL-1998; 98US-0092371.  
PR

XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Sheppard PO, Bairdur N, Delsher TA, Bishop PD, Taft DW;  
XX  
XX WPI; 2001-450736/48.  
DR N-PSDB; AAS11991.  
XX  
PT Disintegrin protease zdin1, useful for producing agents for the  
PT development of antithrombotic and anti-migration of tumour cells and  
PT have antiangiogenic activity -  
XX  
XX Claim 5; Column 53-56; 50pp; English.  
XX  
XX  
XX The invention relates to an isolated novel disintegrin protease family  
XX member, zdin1. Disintegrin bind cell surface molecules, including  
XX integrins, on the surface of various cells such as platelets,  
XX fibroblasts, tumour, endothelial, muscle, neuronal, bone and sperm cells.  
XX Disintegrins are unique and potentially useful tools for investigating  
XX cell matrix and cell-cell interactions, apoptosis, neurogenesis,  
XX connective tissue disorders, chondrogenesis, arthritis, tumour  
XX proliferation, ischaemia reperfusion and inflammation. Additionally, they  
XX are useful in the development of antithrombotic and anti-migration of  
XX tumour cells and have antiangiogenic activity. The present sequence  
XX represents human disintegrin, zdin1, the gene for which is located on  
XX chromosome 2q33.  
XX  
SQ Sequence 696 AA;  
Query Match 100.0%; Score 332; DB 22; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVECDITETCTGSGCCPPNLHK 55  
Db 409 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVECDITETCTGSGCCPPNLHK 463  
RESULT 4  
ABB78130 ID ABB78130 standard; Protein; 696 AA.  
XX  
XX ABB78130;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Amino acid sequence of human zdin1.  
XX  
XX Human; zdin1; disintegrin protease; platelet accumulation;  
KM chromosome 2q33; platelet aggregation; proteolysis; apoptosis;  
KM neurogenesis; myogenesis; connective tissue disorder; arthritis;  
KM chondrogenesis; cell adhesion; cell fusion; neuron; myocyte; heart;  
KM brain; spinal cord; skeletal muscle tissue; Alzheimer's disease;  
KM tumour formation; multiple sclerosis; congestive heart failure;  
KM ischaemic reperfusion; intimal hyperplasia; restenosis; enzyme.  
XX  
XX OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FH Domain 164..382  
FT /note= "protease domain"  
FT 383..464  
FT Domain /note= "disintegrin sequence"  
FT 438..449  
FT Region /note= "disintegrin loop"  
FT 445..696  
FT Domain /note= "cysteine-rich domain"  
XX  
XX US2002072102-A1.  
XX  
XX 13-JUN-2002.  
XX  
XX 16-MAR-2001; 2001US-0809790.  
PF

XX 10-JUL-1998; 98US-092371P.  
PR 09-JUL-1999; 98US-0351414.  
XX  
XX  
PA (SHEP/) SHEPPARD P O.  
PA (BAIN/) BAINDR N.  
PA (DEIS/) DEISHER T A.  
PA (BISH/) BISHOP P D.  
XX  
PI Shepard PO, Baidur N, Deisher TA, Bishop PD;  
XX  
XX WPI: 2002-598452/64.  
DR N-PSDB; ABQ78447.  
XX  
XX New disintegrin homolog polypeptide and polynucleotide, useful for  
PT modulating cell-cell interactions and diagnosis, treatment of  
PT Alzheimer's disease, tumour formation, multiple sclerosis and congestive  
PT heart failure -  
XX  
XX  
XX Claim 7; Page 29-30; 53pp; English.  
XX  
XX The present sequence represents a human polypeptide designated zdintl.  
CC zdintl is a member of the disintegrin protease family. Zdintl inhibits  
CC platelet accumulation. The zdintl gene is present on chromosome 2q33.  
CC Zdintl polypeptides and polynucleotides are useful in treatment of  
CC disorders associated with infarct in brain or heart tissue and/or  
CC platelet aggregation, to modulate proteolysis, apoptosis, neurogenesis,  
CC myogenesis, connective tissue disorders, arthritis, chondrogenesis,  
CC cell adhesion, cell fusion, and signalling or to treat or prevent  
CC development of pathological conditions in such diverse tissue as heart,  
CC brain, spinal cord and skeletal muscle. The molecules modulate  
CC inhibition and proliferation of neurons and myocytes in heart, brain,  
CC spinal cord and skeletal muscle tissue. Disorders which may be amenable  
CC to diagnosis, treatment or prevention with zdintl polypeptides include,  
CC Alzheimer's disease, tumour formation, multiple sclerosis, congestive  
CC heart failure, ischaemic reperfusion or infarct and degenerative  
CC diseases. The zdintl molecules particularly useful in the treatment of  
CC intimal hyperplasia or restenosis due to acute vascular injury.  
XX  
XX  
SQ Sequence 696 AA;  
Query Match 100.0%; Score 332; DB 23; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CSLSGAHCSDGPPCCNNTSCLFQPRGYECRDVAVNECDITTEYCTGDSGQCPNLRK 55  
DB 409 CSLSGAHCSDGPPCCNNTSCLFQPRGYECRDVAVNECDITTEYCTGDSGQCPNLRK 463  
RESULT 5  
AAV30207  
ID AAV30207 standard; Protein; 832 AA.  
XX  
XX AAV30207;  
AC  
XX  
XX 29-OCT-1999 (first entry)  
DT  
XX  
XX Amino acid sequence of the human SVPH3-17 protein.  
DE  
XX  
XX SVPH3-17; metalloproteinase-disintegrin; human chromosome 2;  
KW human chromosome 7; malignant hyperthermia susceptibility;  
KW Zellweger syndrome; neonatal adrenoleukodystrophy;  
KW infantile Refsum disease; progressive familial intraparietal colichestatis;  
KW mucopolysaccharidosis VII; split hand/foot malformation;  
KW arylsulphatase B; right ventricular dysplasia-4; Coppock-like cataract;  
KW insulin dependent diabetes mellitus-12; lamellar type 1 chnhyosis;  
KW transient neonatal myasthenia gravis; congenital aculeiform cataract;  
KW juvenile amyotrophic lateral sclerosis;  
KW familial paroxysmal choreoathetosis;  
KW Finnish lethal neonatal metabolic syndrome.  
XX  
XX Homo sapiens.  
OS

XX  
XX W09941388-A2.  
PN  
XX  
XX 19-AUG-1999.  
PD  
XX  
XX 11-FEB-1999; 99WO-US03016.  
PF  
XX  
XX 11-FEB-1998; 98US-0074310.  
PR  
XX  
XX (IMMV ) IMMUNEX CORP.  
PA  
XX  
XX Cerretti DP;  
PI  
XX  
XX WPI: 1999-527371/44.  
DR N-PSDB; AAZ10207.  
XX  
XX DNA encoding the SVPH3-13 and SVPH3-17 proteins for detecting  
PT disease corresponding to chromosome 7, e.g. Zellweger syndrome  
PT  
XX  
XX  
XX Claim 1; Page 8-9; 82pp; English.  
XX  
XX The present sequence represents a protein designated SVPH3-17, which  
CC is a member of the metalloproteinase-disintegrin family. The  
CC specification also describes SVPH3-13 proteins. Both sequences can  
CC be used to identify human chromosome 2 or 7, and to map genes on these  
CC two chromosomes, and also to identify genes associated with certain  
CC diseases, syndromes, or other human conditions associated with human  
CC chromosome 2 or 7. The disease that correspond to chromosome 7 include  
CC malignant hyperthermia susceptibility, Zellweger syndrome, neonatal  
CC adrenoleukodystrophy, infantile Refsum disease, progressive familial  
CC intraparietal colichestatis, mucopolysaccharidosis VII, and split  
CC hand/foot malformation. Diseases associated with chromosome 2 include  
CC arylsulphatase B; right ventricular dysplasia-4, insulin dependent diabetes  
CC mellitus-12, transient neonatal myasthenia gravis, juvenile amyotrophic  
CC lateral sclerosis, congenital aculeiform cataract, Coppock-like cataract,  
CC lamellar type 1 chnhyosis, familial paroxysmal choreoathetosis, and  
CC Finnish lethal neonatal metabolic syndrome.  
XX  
XX  
SQ Sequence 832 AA;  
Query Match 100.0%; Score 332; DB 20; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7.8e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CSLSGAHCSDGPPCCNNTSCLFQPRGYECRDVAVNECDITTEYCTGDSGQCPNLRK 55  
DB 532 CSLSGAHCSDGPPCCNNTSCLFQPRGYECRDVAVNECDITTEYCTGDSGQCPNLRK 586  
RESULT 6  
AAV25120  
ID AAV25120 standard; Protein; 832 AA.  
XX  
XX AAV25120;  
AC  
XX  
XX 26-AUG-1999 (first entry)  
DT  
XX  
XX Human MDC3 protein.  
DE  
XX  
XX Metalloproteinase-like-disintegrin-like cysteine rich protein; human;  
KW MDC3-alpha; MDC3-beta; MDC3; medical; treatment; diagnosis.  
KW  
OS Homo sapiens.  
XX  
XX JP1115574-A.  
PN  
XX  
XX 15-JUN-1999.  
PD  
XX  
XX 01-DEC-1997; 97JP-0330020.  
PF  
XX  
XX 01-DEC-1997; 97JP-0330020.  
PR  
XX  
XX (EISA ) EISAI CO LTD.  
PA

XX WPI: 1999-398071/34.  
DR N-PSDB; AAX78438.  
XX  
XX  
PT New protein belonging to MDC gene family - useful in medical  
treatment and diagnosis  
XX  
XX Disclosure; Page 13-16; 17pp; Japanese.  
XX  
XX This invention describes novel human MDC2-alpha, MDC2-beta and MDC3  
CC proteins and their encoding nucleic acids. The products of the invention  
CC are useful in medical treatment and diagnosis.  
XX  
SQ Sequence 832 AA;  
  
Query Match 100.0%; Score 332; DB 20; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7,8e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CSLSNGAHGSDGPCNNNTSCLFQPRGYECRDVANECDITETCTGSGQCPNIAHK 55  
Db 532 CSLSNGAHGSDGPCNNNTSCLFQPRGYECRDVANECDITETCTGSGQCPNIAHK 586  
  
RESULT 7  
AAB47778  
ID AAB47778 standard; Protein: 832 AA.  
XX  
XX AAB47778;  
XX  
DT 04-MAR-2002 (first entry)  
XX  
XX ADAM 23.  
XX  
XX Disintegrin; ADAM 23; metalloproteinase; domain; adhesion; proteolysis;  
KM modulator; alpha-v-beta3 integrin; tumour progression; neural tissue;  
KW angiogenesis.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FT Domain 563..570  
FT /label= Disintegrin domain  
XX  
XX MO200174857-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 02-APR-2001; 2001MO-US10729.  
XX  
XX 03-APR-2000; 2000US-194164P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX PI Lopez-Otin C, Freijl JMP, Bianchi AB, Miguel SC, Garcia JML;  
PI Trail P;  
XX  
XX WPI: 2002-066298/09.  
XX  
XX N-PSDB; AAT72025.  
XX  
XX Nucleic acid encoding a new ADAM family member, designated ADAM 23 is  
PT useful to find modulators of its interaction with integrin which can be  
used to prevent angiogenesis or increase neural growth -  
XX  
XX Disclosure; Page 38-41; 44pp; English.  
XX  
XX This sequence shows a cellular disintegrin, ADAM 23, which is a  
CC protein having a disintegrin and metalloproteinase domain. This  
CC protein can perform both adhesion and proteolysis functions.  
CC Modulators of the interaction between ADAM and alpha-v-beta3 integrin  
CC are used to inhibit tumour progression or induce growth of neural  
CC tissue. These compounds modulate angiogenesis and induction of matrix  
CC metalloproteinases facilitating migration of tumour cells and growth of

CC neural tissue.  
XX  
SQ Sequence 832 AA;  
  
Query Match 100.0%; Score 332; DB 23; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7,8e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CSLSNGAHGSDGPCNNNTSCLFQPRGYECRDVANECDITETCTGSGQCPNIAHK 55  
Db 532 CSLSNGAHGSDGPCNNNTSCLFQPRGYECRDVANECDITETCTGSGQCPNIAHK 586  
  
RESULT 8  
ABU56479  
ID ABU56479 standard; Protein: 832 AA.  
XX  
XX ABU56479;  
XX  
DT 02-APR-2003 (first entry)  
XX  
XX Lung cancer-associated polypeptide #72.  
XX  
XX Lung cancer-associated polypeptide; cytosstatic; emphysema;  
KW antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
XX  
XX Unidentified.  
XX  
XX WO200286443-A2.  
XX  
XX 31-OCT-2002.  
XX  
XX 18-APR-2002; 2002MO-US12476.  
XX  
XX 18-APR-2001; 2001US-284770P.  
XX  
XX 10-MAY-2001; 2001US-290492P.  
XX  
XX 09-NOV-2001; 2001US-339245P.  
XX  
XX 13-NOV-2001; 2001US-350666P.  
XX  
XX 29-NOV-2001; 2001US-334370P.  
XX  
XX 12-APR-2002; 2002US-372246P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
XX Aziz N, Murray R;  
XX  
XX WPI: 2003-093161/08.  
XX  
XX N-PSDB; ABX76201.  
XX  
XX  
XX Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
XX expression in lung cancer -  
XX  
XX  
XX Claim 27; Page 243-244; 453pp; English.  
XX  
XX The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung  
CC cancer-associated polynucleotides and polypeptides are used for  
CC identifying a compound that modulates a lung cancer-associated  
CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
CC cell to treat lung cancer in a patient and for treating a mammal having  
CC lung cancer by administering a modulatory compound identified. The  
CC methods are useful for treating lung cancer, such as small cell lung  
CC cancer, non-small cell lung cancer or other benign or precancerous  
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides

CC and polypeptides are useful for diagnostic purposes and as targets for  
CC screening for therapeutic compounds that modulate lung cancer, such as  
CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated  
CC polypeptides of the invention.

XX Sequence 832 AA;

Query Match 100.0%; Score 332; DB 24; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7.8e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVANECDITTEYCTGDSGQCPNHLK 55  
Db 532 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVANECDITTEYCTGDSGQCPNHLK 586

RESULT 9

ABUS6563  
ID ABUS6563 standard; Protein; 832 AA.

XX ABUS6563;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #156.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;

KW antiinflammatory; antisthmatic; non-small cell lung cancer; atelectasis;

KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO00286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US12476.

XX 18-APR-2001; 2001US-284770P.

XX 10-MAY-2001; 2001US-290492P.

XX 09-NOV-2001; 2001US-339245P.

XX 13-NOV-2001; 2001US-350666P.

XX 29-NOV-2001; 2001US-34370P.

XX 12-APR-2002; 2002US-372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX N-PSDB; ABX76292.

XX The invention relates to a method for detecting a lung cancer-associated

CC transcrip in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises

CC to a sequence that is at least 80 % identical to a gene that exhibits

CC increased or decreased expression in lung cancer samples. Lung

CC cancer-associated polynucleotides and polypeptides are used for

CC identifying a compound that modulates a lung cancer-associated

CC polypeptide, for inhibiting proliferation of a lung cancer-associated

CC cell to treat lung cancer in a patient and for treating a mammal having

CC lung cancer by administering a modulatory compound identified. The

CC methods are useful for treating lung cancer, such as small cell lung

CC cancer, non-small cell lung cancer or other benign or precancerous

CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
CC and polypeptides are useful for diagnostic purposes and as targets for  
CC screening for therapeutic compounds that modulate lung cancer, such as  
CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated  
CC polypeptides of the invention.

XX Sequence 832 AA;

Query Match 100.0%; Score 332; DB 24; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7.8e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVANECDITTEYCTGDSGQCPNHLK 55  
Db 532 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVANECDITTEYCTGDSGQCPNHLK 586

RESULT 10

AAR67757  
ID AAR67757 standard; Protein; 488 AA.

XX AAR67757;

DT 25-MAR-2003 (updated)

DT 01-AUG-1995 (first entry)

DE Human fetal brain MDC protein.

XX MDC protein; breast cancer; mamma carcinoma; ovary cancer;

KW chromosome-17.

XX Homo sapiens.

XX EP633268-A2.

XX 11-JAN-1995.

XX 13-MAY-1994; 94EP-0107487.

XX 14-MAY-1993; 93JP-0136602.

XX 22-SEP-1993; 93JP-0257455.

XX 23-FEB-1994; 94JP-0049904.

XX 12-APR-1994; 94JP-0073328.

XX (CANC-) CANCER INST.

XX (ETISA ) EISAI CO LTD.

XX Emi M, Nakamura Y;

XX WPI; 1995-038478/06.

XX N-PSDB; AAQ76119.

XX Novel MDC protein and DNA encoding it - used to develop prods.

XX for the study, diagnosis and therapy of cancers, partic. breast

XX and ovarian cancer

XX Disclosure, Page 73-74, 123pp; English.

XX A detailed map of human chromosome-17 was constructed to analyze the

CC chromosome in breast and ovarian cancer tissues. 2 Novel cDNAs

CC (given in AAQ76120-21) were obtained from fetal cerebellum that

CC encoded novel MDC proteins (AAR67758-59, respectively). 5'-RACE and

CC RT-PCR revealed the sequence given in AAQ76122, encoding the MDC given

CC in AAR7552. Sequences common for these clones are given in AAQ76119

CC and AAR6757, respectively. A genomic DNA sequence (AAQ76124) for MDC

CC was obtained from cosmid clone CC117-904.

XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 488 AA;

Query Match 64.2%; Score 213; DB 16; Length 488;

Mon Oct 27 15:28:06 2003

us-09-634-252a-4\_copy\_532\_586.rag

Page 7

Best Local Similarity 65.5%; Pred. No. 2.9e-13;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Qy 1 CSLSNAGHCSDBGPCNTSCIFQPRGIECRDAVNECDITEYCTGSGGQCPNHLK 55  
 Db 371 CTLTHDAMCSDGLCCRR--CKYEPRGVSCREAVNECDIAETCTGDSQCPNHLK 423

```
Search completed: October 21, 2003, 10:08:28
Job time : 38.0126 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 09:02:08 ; Search time 9.68553 Seconds  
(without alignments)  
267.045 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGAHCSGDCPCNNNTSC.....CDITEXCTGDI;GQCPPLMK 55

Scoring table: BLOSUM62  
Gapop 10.0 , Gapexc 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	213	64.2	769	1	AD11_HUMAN	O75078 homo sapien
2	213	64.2	773	1	AD11_MOUSE	O91144 mus musculu
3	185	55.7	452	1	AD11_XENLA	O90523 xenopus lae
4	185	55.7	906	1	AD22_HUMAN	O90081 homo sapien
5	184	55.4	857	1	AD22_MOUSE	O91144 mus musculu
6	183	55.1	935	1	AD22_XENLA	O42396 xenopus lae
7	173.5	52.3	816	1	AD15_RAT	O91144 mus musculu
8	168.5	50.8	814	1	AD15_HUMAN	O91144 mus musculu
9	167	50.3	790	1	AD30_HUMAN	O91144 mus musculu
10	160	48.2	813	1	AD33_HUMAN	O91144 mus musculu
11	158.5	47.7	815	1	AD15_MOUSE	O91144 mus musculu
12	153	46.1	735	1	AD02_CAYPO	O60411 cavia porce
13	152	45.8	735	1	AD02_HUMAN	O91144 mus musculu
14	152	45.8	751	1	AD02_RABIT	O28660 oryctolagus
15	151	45.5	735	1	AD02_MACFA	O28478 macaca fasc
16	147	44.3	616	1	ECAR_ECHCA	O90495 echis carin
17	145.5	43.8	571	1	DIST_BORTA	O30431 bothrops ja
18	145	43.7	722	1	AD21_HUMAN	O91144 mus musculu
19	144	43.4	735	1	AD02_MOUSE	O91144 mus musculu
20	143	43.1	697	1	AD26_MOUSE	O91144 mus musculu
21	143	43.1	820	1	AD29_HUMAN	O91144 mus musculu
22	143	43.1	920	1	AD19_HUMAN	O91144 mus musculu
23	142	42.8	909	1	AD12_HUMAN	O91144 mus musculu
24	141	42.5	797	1	AD33_MOUSE	O91144 mus musculu
25	140.5	42.3	761	1	AD24_MOUSE	O91144 mus musculu
26	140	42.2	726	1	AD20_HUMAN	O91144 mus musculu
27	138	41.6	760	1	AD25_MOUSE	O91144 mus musculu
28	137.5	41.4	737	1	AD02_RAT	O91144 mus musculu
29	134	40.4	416	1	HRIB_TRIPL	O77780 bos taurus
30	134	40.4	745	1	AD02_BOVIN	O91144 mus musculu
31	134	40.4	903	1	AD12_MOUSE	O91144 mus musculu
32	132.5	39.9	824	1	AD17_HUMAN	O91144 mus musculu
33	132	39.8	789	1	AD07_RAT	O91180 rattus norv

## ALIGNMENTS

34	132	39.8	956	1	AD19_HUMAN	O90113 homo sapien
35	130.5	39.3	827	1	AD17_MOUSE	O92018 mus musculu
36	130.5	39.3	827	1	AD17_RAT	O92169 rattus norv
37	130	39.2	776	1	AD28_MACFA	O91436 macaca fasc
38	130	39.2	819	1	AD09_HUMAN	O13443 homo sapien
39	129	38.9	729	1	AD21_MOUSE	O91144 mus musculu
40	129	38.9	788	1	AD07_MOUSE	O35227 mus musculu
41	127	38.3	719	1	AD18_MOUSE	O91157 mus musculu
42	126	38.0	445	1	AD18_RAT	O91144 mus musculu
43	125	38.0	739	1	AD18_HUMAN	O91144 mus musculu
44	125.5	37.8	65	1	DISE_CERCE	P83041 cerastes ce
45	125	37.7	746	1	AD18_MACFA	O91194 macaca fasc

RESULT 1  
AD11\_HUMAN STANDARD; PRT; 769 AA.  
AC O75078: Q14808: Q14809: Q14810:  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE ADAM 11 precursor (A disintegrin-like and metalloproteinase domain 11)  
DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)  
DE (MDC).  
GN ADAM11 OR MDC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC TISSUE=Brain;  
RX MEDLINE=98359734; Pubmed=9693107;  
RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.;  
RT "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2  
RT and MDC3: novel human cellular disintegrins highly expressed in the  
RT brain";  
RL Biochem. J. 334:93-98(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC TISSUE=Cerebellum;  
RX MEDLINE=94073190; Pubmed=8252040;  
RA Kasumi F., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I.,  
RT "Human metalloproteinase/disintegrin-like gene at 17q21.3 is  
RT somatically rearranged in two primary breast cancers";  
RL Nat. Genet. 5:151-157(1993).  
RN [3]  
RP SEQUENCE OF 106-769 FROM N.A. (ISOFORMS LONG AND SHORT).  
RC TISSUE=Brain, Breast, Ovary, and Testis;  
RX MEDLINE=95044425; Pubmed=7956356;  
RA Katagiri T., Harada Y., Emi M., Nakamura Y.;  
RT "Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron  
RT organization and alternative splicing";  
RL Cytogenet. Cell Genet. 68:39-44(1995).  
RN [4]  
RP FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON  
CC CATALYTIC METALLOPROTEIN.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long; Synonyms=MDC-769;  
CC Name=Short; Synonyms=MDC-524;  
CC IsoId=O75078-2; Sequence=VSP\_005472, VSP\_005473, VSP\_005474,  
CC VSP\_005475;  
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN. SLIGHTLY  
CC DETECTED OR NOT AT ALL IN OTHER TISSUES.  
CC -1- DOMAIN: A CONSERVED MOTIF (AVN(E/D)CD) WITHIN THE DISINTEGRIN-LIKE  
CC DOMAIN COULD BE INVOLVED IN THE BINDING TO THE INTEGRIN RECEPTOR.  
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY

CC	SIMILARITY).
CC	- - DISEASE : REARRANGEMENTS OCCUR IN BREAST AND OVARIAN CANCERS WHICH INVOLVE MULTIPLE EXONS AND DISRUPT THE CODING REGION.
CC	- - SIMILARITY : Belongs to peptidase family M12B.
CC	- - SIMILARITY : Contains 1 EGF-like domain.
CC	- - SIMILARITY : Contains 1 disintegrin domain.
CC	-----
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; AB009675; BAA3352.1; ..
DR	EMBL; D17390; BAA04213.1; ..
DR	EMBL; D31872; BAA06670.1; ..
DR	EMBL; D31872; BAA06671.1; ..
DR	PIR; I65967; I65967.
DR	HSSP; P18619; 1FVL.
DR	MEROPS; M12_976; ..
DR	Genew; HGNC:189; ADAM11.
DR	MIM; 155120; ..
DR	GO; GO:0005886; C:plasma membrane; TAS.
DR	GO; GO:0005178; F:integrin binding activity; TAS.
DR	GO; GO:0007228; F:integrin-mediated signaling pathway; TAS.
DR	InterPro; IPRO06586; ADAM_cysteine.
DR	InterPro; IPRO01762; Disintegrin.
DR	InterPro; IPRO06209; EGF like.
DR	InterPro; IPRO02870; Pep_M12B_propep.
DR	InterPro; IPRO01590; Reprolysin.
DR	InterPro; IPRO06025; Zn_M12peptidase.
DR	Pfam; PF00200; disintegrin; 1.
DR	Pfam; PF01562; Pep_M12B_propep; 1.
DR	Pfam; PF01421; Reprolysin; 1.
DR	PRINTS; PR00289; DISINTEGRIN.
DR	ProDom; PD000664; Disintegrin; 1.
DR	SMART; SMO0608; ACR; 1.
DR	SMART; SMO0050; DISIN; 1.
DR	PROSITE; PS50215; ADAM_MERPO; 1.
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR	PROSITE; PS50214; DISINTEGRIN_2; 1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; FALSE_NEG.
DR	PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW	Signal; Glycoprotein; Transmembrane; EGF-like domain; Alternative splicing.
FT	SIGNAL 1 23
FT	PROPEP 24 225
FT	CHAIN 226 769
FT	DOMAIN 226 734
FT	TRANSMEM 735 755
FT	DOMAIN 756 769
FT	DOMAIN 226 438
FT	DOMAIN 444 511
FT	DOMAIN 532 676
FT	DOMAIN 677 709
FT	DISULFID 349 433
FT	DISULFID 503 516
FT	DISULFID 677 692
FT	DISULFID 686 698
FT	DISULFID 700 709
FT	CARBOHYD 96 163
FT	CARBOHYD 605 605
FT	CARBOHYD 673 673
FT	VARSPLIC 1 99
FT	VARSPLIC 100 104
FT	FT/FLN-VSP 005473.
FT	DIELN -> MCWL5 (in isoform Short).
FT	/FTLVSP 005473
FT	DYLCPLQCNISGAPRIADLVGDLSVLT -> PQGGAVW
FT	LPPICQLHWSSSARGGGGHQ (in isoform Short).

```

FT FT VASSPLIC 624 769 /FtId=VSP_005474. Missing (in isoform Short).
FT FT CONFLICT 106 106 /FtId=VSP_005475.
FT FT CONFLICT 325 325 Q -> H (IN REF. 1) .
SQ SEQUENCE 769 AA; 83409 MW; 59B9C457B6DDFB6E CRC64; D -> N (IN REF. 2) .

Query Match 64.2%; Score 213; DB 1; Length 769;
Best Local Similarity 65.5%; Pred. No. 3,1e-15;
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Cy 1 CSLNSGACSHCDGPCCNNNTSCIFQPGREYECRDVAVECDITTECTGTGDSGCQEPNLHK 55
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 477 CTHLHDAMCSGLCCRR--CKYERGRYSCHRAVNECIAETICTDSSQCPEPNLHK 529

RESULT 2
AD1_MOUSE STANDARD; PRT; 773 AA.
ID AD1_MOUSE
AC Q9RLV4, 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last modification update)
DE ADAM 1 precursor (A disintegrin and metalloproteinase domain 1)
DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)
DE (MDC).
GN ADAM1 OR MDC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99365303; PubMed=10433968;
RA Sagane K., Yamazaki K., Mizui Y., Tanaka I.;
RT "Cloning and chromosomal mapping of mouse ADAM1, ADAM2 and ADAM3.";
RL Gene 236.79-86(1999).
CC - FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON
CC CATALYTIC METALLOPROTEASE-LIKE PROTEIN.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE BRAIN. WEAKLY
CC DETECTED IN THE HEART, LIVER AND TESTIS.
CC - DOMAIN: A CONSERVED MOTIF [AVN(E/D)CD] WITHIN THE DISINTEGRIN-LIKE
CC DOMAIN COULD BE INVOLVED IN THE BINDING TO THE INTEGRIN RECEPTOR.
CC - PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC - SIMILARITY: Belongs to peptidase family M12B.
CC - SIMILARITY: Contains 1 EGF-like domain.
CC - SIMILARITY: Contains 1 disintegrin domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/news/e/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB009676; BAA83384.1; -
DR HSBP; P18619; IPVL.
DR MEROPS; M12_976; -.
DR MGJ; MGJ:1098667; Adam1.
DR InterPro; IPR006586; ADM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR006025; Zn_MTpeptide.
DR Pfam; PF00200; disintegrin_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Regiolysin_1.
DR PRINTS; PR00289; DISINTEGRIN.

```

DR Prodom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PSS0215; ADAM\_MERPRO; 1; FALSE\_NEG.  
 DR PROSITE; PSS0214; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PSS0214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PSS0022; EGF\_1; 1.  
 DR PROSITE; PSS0186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; FALSE\_NEG.  
 DR Signal; Glycoprotein; Transmembrane; EGF-like domain.  
 KM SIGNAL; Glycoprotein; Transmembrane; EGF-like domain.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT PROPEP 25 229 BY SIMILARITY.  
 FT CHAIN 230 773 ADAM 11.  
 FT DOMAIN 230 738 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 739 759 POTENTIAL.  
 FT DOMAIN 760 773 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 230 442 METALLOPROTEASE-LIKE.  
 FT DOMAIN 448 535 DISINTEGRIN-LIKE.  
 FT DOMAIN 536 680 CYS-RICH.  
 FT DOMAIN 681 713 EGF-LIKE.  
 FT DISULFID 353 437 BY SIMILARITY.  
 FT DISULFID 507 520 POTENTIAL.  
 FT DISULFID 681 696 BY SIMILARITY.  
 FT DISULFID 690 702 BY SIMILARITY.  
 FT DISULFID 704 713 BY SIMILARITY.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 773 AA; 84134 MW; 9A8A125FB3F743DD CRC64;

Query Match 64.2%; Score 213; DB 1; Length 773;  
 Best Local Similarity 65.5%; Pred. No. 3; 1e-15;  
 Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;  
 Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDVAVECDITEYCTGDSGQCPNHLK 55  
 Db 481 CTLHDMCMCDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGQCPNHLK 533

RESULT 3  
 AD11\_XENLA STANDARD; PRT; 452 AA.  
 AC O9PS23;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE ADAM 11 (A disintegrin and metalloproteinase domain 11)  
 DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)  
 DE (MDC) (Metalloproteinase-disintegrin MDC11a) (MDC11.1) (Fragment).  
 GN ADAM11 OR MDC11a.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=99102794; PubMed=9882486;  
 RA Cal H., Kraetzschmar J., Alfandari D., Hunnicutt G., Blobel C.P.;  
 RT "Neural crest-specific and general expression of distinct  
 RT metalloproteinase-disintegrins in early Xenopus laevis development.";   
 RL Dev. Biol. 204:508-524(1998).  
 CC -1- FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON  
 CC CATALYTIC METALLOPROTEASE-LIKE PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: DETECTED IN TESTIS AND BARELY EXPRESSED IN  
 CC HEART AND MUSCLE. NOT DETECTABLE IN LIVER.  
 CC -1- DEVELOPMENTAL STAGE: COULD NOT BE DETECTED IN EMBRYOS UNTIL  
 CC NEURULATION. IN DEVELOPING EMBRYOS, THE EXPRESSION IS RESTRICTED  
 CC TO NEURAL CREST DERIVATIVES.  
 CC -1- DOMAIN: A CONSERVED MOTIF [AVN(E/D)CD] WITHIN THE DISINTEGRIN-LIKE

CC DOMAIN COULD BE INVOLVED IN THE BINDING TO THE INTEGRIN RECEPTOR.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF032384; AAC61848.1; -  
 DR HSPSP; P18619; 1FVL.  
 DR MEROPS; M12; UNB; -  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR001590; Reprolysin.  
 DR InterPro; IPR006025; Zn\_MTPeptide.  
 DR Pfam; PF00200; disintegrin; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR Prodom; PD000664; Disintegrin; 1.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PSS0215; ADAM\_MERPRO; 1.  
 DR PROSITE; PSS0427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PSS0214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PSS0022; EGF\_1; 1.  
 DR PROSITE; PSS0186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; FALSE\_NEG.  
 KM Glycoprotein; Transmembrane; EGF-like domain.  
 FT NON TER 1 1  
 FT DOMAIN 1 417 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 418 438 POTENTIAL.  
 FT DOMAIN 439 452 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 126 214 METALLOPROTEASE-LIKE.  
 FT DOMAIN 215 359 DISINTEGRIN-LIKE.  
 FT DOMAIN 360 416 CYS-RICH.  
 FT DISULFID 31 115 EGF-LIKE.  
 FT DISULFID 186 199 BY SIMILARITY.  
 FT DISULFID 360 375 POTENTIAL.  
 FT DISULFID 369 381 BY SIMILARITY.  
 FT DISULFID 383 392 BY SIMILARITY.  
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 452 AA; 48577 MW; DF1F66C24CCD6847 CRC64;  
 Query Match 55.7%; Score 185; DB 1; Length 452;  
 Best Local Similarity 56.4%; Pred. No. 1; 7e-12;  
 Matches 31; Conservative 8; Mismatches 14; Indels 2; Gaps 1;  
 Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDVAVECDITEYCTGDSGQCPNHLK 55  
 Db 160 CTLHDMCMCDGLCCRR--GCKYEPBGVSCREAVNECDVPEACPDSSACPNHLK 212  
 RESULT 4  
 AD22\_HUMAN STANDARD; PRT; 906 AA.  
 AC O9P0K1; O75075; O75076; O9P0K2; O9U1A1; O9U1K2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAM 22 precursor (A disintegrin and metalloproteinase domain 22)

DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein  
DE 2) (Metalloproteinase-disintegrin ADAM22-3).  
CN ADAM22 OR MDC2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=20504287; PubMed=11050470;  
RA Harada T., Nishie A., Torigoe K., Ikeraki K., Shono T., Maehara Y.,  
RA Kuwano M., Wada M.,  
RT "The specific expression of three novel splice variant forms of human  
RT metalloproteinase-like disintegrin-like cysteine-rich protein 2 gene  
RT in brain tissues and gliomas."  
RN Jpn. J. Cancer Res. 91:1001-1006(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
RC TISSUE=Brain;  
RX MEDLINE=98359734; PubMed=9693107;  
RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.,  
RT "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2  
RT and MDC3: novel human cellular disintegrins highly expressed in the  
RT brain."  
RN Biochem. J. 334:93-98(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 5).  
RC TISSUE=Brain;  
RA Wada M., Torigoe K., Harada T., Kuwano M.,  
RT "Isolation and tissue specific expression of novel ADAM family from  
RT 7q21.1 region."  
RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 40-906 FROM N.A. (ISOFORM 1).  
RC TISSUE=Cerebellum;  
RX MEDLINE=99453762; PubMed=10524237;  
RA Poindecker K., Nelson N., Dubose R.F., Black R.A., Cerretti D.P.,  
RT "The identification of seven metalloproteinase-disintegrin (ADAM)  
RT genes from genomic libraries."  
RN Gene 237:61-70(1999).  
CC -1- FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON  
CC CATALYTIC METALLOPROTEINASE-LIKE PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=5;  
CC Name=1; Synonyms=Epsilon;  
CC IsoId=Q9P0K1-1; Sequence=Displayed;  
CC Name=2; Synonyms=Delta;  
CC IsoId=Q9P0K1-2; Sequence=VSP\_005482, VSP\_005484;  
CC Name=3; Synonyms=Alpha;  
CC IsoId=Q9P0K1-3; Sequence=VSP\_005483;  
CC Name=4; Synonyms=Beta;  
CC IsoId=Q9P0K1-4; Sequence=VSP\_005482, VSP\_005483;  
CC Name=5;  
CC IsoId=Q9P0K1-5; Sequence=VSP\_005482;  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE BRAIN. DETECTED SLIGHTLY  
CC OR NOT AT ALL IN OTHER TISSUES.  
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 disintegrin domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF155382; AAF73289.1; -;  
CC EMBL; AF155381; AAF73288.1; -;

DR EMBL; AB009671; BAA32349.1; -;  
DR EMBL; AB009671; BAA32350.1; -;  
DR EMBL; AF073291; AAF22476.2; -;  
DR EMBL; AF158637; AAD55251.1; -;  
DR HSRP; P18619; 1FVL.  
DR MEROPS; M12.978; -;  
DR MIM; 603709; -;  
DR GeneW; HGNC:201; ADAM22.  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0005178; F:integrin binding activity; NAS.  
DR GO; GO:0007162; P:negative regulation of cell adhesion; NAS.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001762; Disintegrin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR002870; Pcp\_M12B\_propep.  
DR InterPro; IPR001590; Reprolysin.  
DR InterPro; IPR006025; Zn\_MTPeptide.  
DR Pfam; PF01562; Pcp\_M12B\_propep; 1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR PRINTS; PR00289; Disintegrin.  
DR PRODOM; PD000664; Disintegrin; 1.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00050; D1SIN; 1.  
DR PROSITE; PS50215; ADAM\_MEROP; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
DR Signal; GlycoProtein; Transmembrane; EGF-like domain;  
KW Alternative splicing.  
FT SIGNAL 1 25  
FT PROPEP 26 222  
FT CHAIN 223 906  
FT DOMAIN 223 736  
FT TRANSMEM 737 757  
FT DOMAIN 758 906  
FT DOMAIN 223 438  
FT DOMAIN 444 531  
FT DOMAIN 532 678  
FT DOMAIN 675 712  
FT DISULFID 349 433  
FT DISULFID 503 516  
FT DISULFID 679 694  
FT DISULFID 688 700  
FT DISULFID 702 711  
FT CARBOHYD 175 175  
FT CARBOHYD 519 519  
FT CARBOHYD 634 634  
FT CARBOHYD 675 675  
FT VARSPPLIC 768 803  
FT /FTId=VSP\_005482.  
FT /FTId=VSP\_005482.  
FT VARSPPLIC 860 906  
FT /FTId=VSP\_005483.  
FT VARSPPLIC 859 859  
FT /FTId=VSP\_005484.  
FT CONFLICT 81 81  
FT SEQUENCE 906 AA; 100432 MW; 265ECCDFAC088B CRC64;  
SQ

Query Match 55.7%; Score 185; DB 1; Length 906;  
Best Local Similarity 56.4%; Pred. No. 3e-12;  
Matches 31; Conservative 8; Mismatches 14; Indels 2; Gaps 1;

Qy 1 CSLSGAHGSDGPPCCNNTSLCPQPGYECRDVAVECDITTEYCTGDSGCPNPLHK 55  
Db 477 CTLTGDSQSCDGLCK--KCKFQPGYTCREAVNDCCDIRFCGNSGCAINIK 529

RESULT 5  
AD22\_MOUSE

ID	AD22 MOUSE	STANDARD:	PRT:	857 AA.
AC	Q9R1V6; Q9R1V5;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	ADAM 22 precursor (A disintegrin and metalloproteinase domain 22).			
GN	ADAM22.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).			
RP	TISSUE=Brain;			
RC	MEDLINE=99365303; PubMed=10433968;			
RX	Sagane K., Yamazaki K., Mizui Y., Tanaka I.;			
RA	"Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23,"			
RL	Gene 236:79-86(1999).			
RT				
CC	- FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON			
CC	CATALYTIC METALLOPROTEASE-LIKE PROTEIN.			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=2;			
CC	Name=Alpha;			
CC	IsoId=Q9R1V6-1; Sequence=Displayed;			
CC	Name=Beta;			
CC	IsoId=Q9R1V6-2; Sequence=VSP_005485;			
CC	- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE BRAIN.			
CC	- PTR: THE PRECURSOR IS CLEAVED BY A FORIN ENDOPEPTIDASE (BY			
CC	SIMILARITY).			
CC	- SIMILARITY: Belongs to peptidase family M12B.			
CC	- SIMILARITY: Contains 1 EGF-like domain.			
CC	- SIMILARITY: Contains 1 disintegrin domain.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC				
DR	EMBL; AB009674; BAAG3383.1; -			
DR	EMBL; AB009674; BAA83383.1; -			
DR	HSSP; P18619; 1FVL.			
DR	MEROPS; M12.978; -			
DR	MGP; MGI:1340046; Adam22.			
DR	InterPro; IPR006586; ADAM_cysteine.			
DR	InterPro; IPR001762; Disintegrin.			
DR	InterPro; IPR006209; EGF-like.			
DR	InterPro; IPR002870; Pep_M12B_propep.			
DR	InterPro; IPR001590; Reprolysin.			
DR	InterPro; IPR006025; Zn_MTPeptide.			
DR	Pfam; PF00200; disintegrin; 1.			
DR	Pfam; PF01562; Pep_M12B_propep; 1.			
DR	Pfam; PF01421; Reprolysin; 1.			
DR	ProDom; PD00664; Disintegrin; 1.			
DR	SMART; SM00608; ACR; 1.			
DR	SMART; SM00050; DISIN; 1.			
DR	PROSITE; PS00215; ADAM_MERPO; 1.			
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.			
DR	PROSITE; PS00214; DISINTEGRIN_2; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS01186; EGF_2; FALSE_NEG.			
DR	PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.			
KW	Signal; Glycoprotein; Transmembrane; EGF-like domain;			
KW	Alternative splicing.			
FT	SIGNAL	1	23	POTENTIAL.
FT	PROPEP	24	223	BY SIMILARITY.
FT	CHAIN	224	857	ADAM 22.
FT	DOMAIN	24	734	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	735	755	POTENTIAL.
FT	DOMAIN	756	857	CYTOPLASMIC (POTENTIAL).

Query Match	Best Local Similarity	Score 184; DB 1; Length 857;	Match	Similarity	56.4%; Pred. No. 3,78-12;	Matches	31; Conservative 8; Pmiatches 14; Indels 2; Gaps 1;
FT	DOMAIN	224	436				
FT	DOMAIN	442	529				
FT	DOMAIN	533	666				
FT	DOMAIN	673	710				
FT	DISULFID	347	431				
FT	DISULFID	501	514				
FT	DISULFID	677	692				
FT	DISULFID	686	698				
FT	DISULFID	700	709				
FT	CARBOHYD	163	163				
FT	CARBOHYD	517	517				
FT	CARBOHYD	632	632				
FT	CARBOHYD	673	673				
FT	VARSPLIC	766	801				
SO	SEQUENCE	857 AA;	94740 MM;				
Qy	1	CSLSNGAHCSGPGCCNTSCIFGRGRCRDVAVNECITVECTDSQCSPNLIK	55	475	CTLTPDSQCSGLCK--KCRFPLGTVCRAVNDICREICGNSSCAPVAK	527	
Db	475	CTLTPDSQCSGLCK--KCRFPLGTVCRAVNDICREICGNSSCAPVAK	527				
AD22_XENLA	STANDARD;	PRT;	935 AA.				
AC	042596;						
DT	16-OCT-2001 (Rel. 40, Created)						
DT	16-OCT-2001 (Rel. 40, Last sequence update)						
DT	16-OCT-2001 (Rel. 40, Last annotation update)						
DE	ADAM 22 precursor (A disintegrin and metalloproteinase domain 22)						
DE	(Metalloprotease-disintegrin MDC1b) (MDC11.2).						
GN	ADAM22 OR MDC1B.						
OS	Xenopus laevis (African clawed frog).						
OC	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Xenopus; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;						
OC	Xenopodinae; Xenopus.						
OX	NCBI_TaxID=8355;						
OX	[1]						
RN	SEQUENCE FROM N.A.						
RP	TISSUE=Testis;						
RC	TISSUE=Testis;						
RC	MEDLINE=97349132; PubMed=9205136;						
RA	Shilling F.M., Kraetzschmar J., Cai H., Weiskamp G., Gayko U.,						
RA	Leibow J., Wyles D.G., Nuccitelli R., Blobel C.P.;						
RT	"Identification of metalloprotease/disintegrin in Xenopus laevis						
RT	testis with a potential role in fertilization.";						
RL	Dev. Biol. 186:155-164(1997).						
RL	- FUNCTION: PROBABLE LIGAND FOR INTEGRIN IN THE BRAIN. THIS IS A NON						
CC	CATALYTIC METALLOPROTEASE-LIKE PROTEIN (BY SIMILARITY).						
CC	- SUBCELLULAR LOCATION: Type I membrane protein.						
CC	- TISSUE SPECIFICITY: LOW LEVELS IN ADULT TISSUES. NOT DETECTED IN						
CC	DEVELOPING EMBRYOS.						
CC	- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY						
CC	SIMILARITY).						
CC	- SIMILARITY: Belongs to peptidase family M12B.						
CC	- SIMILARITY: Contains 1 EGF-like domain.						
CC	- SIMILARITY: Contains 1 disintegrin domain.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on these						
CC	use by non-profit institutions as long as its content is in no way						

modified and this statement is not removed. Usage by and for commercial  
entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL: AF033383; AAC61847.1; -  
DR EMBL: U78188; AAB87148.1; -  
DR HSSP: P17347; 2ECH.  
DR MEMOPS: M12.UBP; -  
DR InterPro: IPR006586; ADAM cysteine.  
DR InterPro: IPR001762; Disintegrin.  
DR InterPro: IPR006209; EGF like.  
DR InterPro: IPR006210; EGF.  
DR InterPro: IPR002870; Peg\_M12B\_propep.  
DR InterPro: IPR001590; ReprOlysin.  
DR InterPro: IPR006025; Zn\_MTPeptide.  
DR Pfam: PF00200; disintegrin; 1.  
DR Pfam: PF01562; Peg\_M12B\_propep; 1.  
DR Pfam: PF01421; ReprOlysin; 1.  
DR ProDom: PD000664; Disintegrin; 1.  
DR SMART: SM00608; ACR; 1.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS50215; ADAM\_MEMPRO; 1.  
DR PROSITE: PS00427; DISINTEGRIN\_1; 1.  
DR PROSITE: PS50214; DISINTEGRIN\_2; 1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE: PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
DR Signal: Glycoprotein; Transmembrane; EGF-like domain.  
FT SIGNAL 1 24  
FT PROPEP 25 227  
FT CHAIN 228 935  
FT DOMAIN 228 736  
FT TRASMEM 737 757  
FT DOMAIN 758 935  
FT DOMAIN 228 440  
FT DOMAIN 446 533  
FT DOMAIN 536 670  
FT DOMAIN 677 713  
FT DISULFID 351 435  
FT DISULFID 505 518  
FT DISULFID 681 695  
FT DISULFID 689 701  
FT DISULFID 703 712  
FT CARBOHYD 167 167  
FT CARBOHYD 210 210  
FT CARBOHYD 521 521  
FT CARBOHYD 609 609  
FT CARBOHYD 636 636  
FT CARBOHYD 677 677  
SQ SEQUENCE 935 AA; 104161 MW; 77B7AFDCC5C77C90 CRC64;

Query Match 55.1%; Score 183; DB 1; Length 935;  
Best Local Similarity 54.5%; Pred. No. 5, 1e-12;  
Matches 30; Conservative 10; Mismatches 13; Indels 2; Gaps 1;

Qy 1 CSLNGAHGSDPCPCNNNTSCLFQPRGYECRDVNECDITEYCTGDSGQCPNLIK 55  
Db 479 CTLTDSBSCSDLCCSN--CKFNPKMCLREAVNDCDIFETCTGNTSGCPAIIHK 531

RESULT 7  
AD15\_RAT  
ID AD15\_RAT STANDARD; PRT; 816 AA.  
AC O90YV0:  
DT 16-OCT-2001 (Rel. 40; Created)  
DT 16-OCT-2001 (Rel. 40; Last sequence update)  
DT 28-FEB-2003 (Rel. 41; Last annotation update)  
DE ADAM\_15 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
domain 15) (Metalloproteinase-like, disintegrin-like, and cysteine-  
rich protein 15) (MDC-15) (Metalloprotease RGD disintegrin protein)  
DE (Metalloprotein) (CRII-7).  
GN ADAM15 OR MDC15.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxId=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Sciatic nerve;  
RX MEDLINE=20556183; PubMed=11102971;  
RA Bosse F., Petzold G., Greiner-Petter R., Pipplirs U., Gillen C.,  
RA Mueller H.-W.;  
RT "Cellular localization of the disintegrin CRII-7/MDC15 mRNA in rat  
RT PNS and CNS and regulated expression in postnatal development and  
RT after nerve injury."  
RT RL Gila 32:313-327(2000).  
CC - FUNCTION: May be involved in cell-surface proteolysis, cell  
CC adhesion or intracellular protein maturation.  
CC - COPACITOR: Binds 1 zinc ion per subunit (Potential).  
CC - SUBUNIT: Interacts with ITGA5-ITB3 (vitronectin receptor),  
CC PACSIN3 and SNX9. PACSIN3 and SNX9 preferentially bind the  
CC precursor but not the processed form of ADAM15, suggesting that  
CC the interaction occurs in a secretory pathway compartment prior to  
CC the medial Golgi (By similarity). Interacts specifically with Src  
CC family protein-tyrosine kinases (PTKs) (By similarity).  
CC - SUBCELLULAR LOCATION: Type I membrane protein.  
CC - TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN, SPINAL CORD,  
CC SCIATIC NERVE AND LUNG. EXPRESSED AT LOWER LEVELS IN ALL OTHER  
CC TISSUES. IN THE PERIPHERAL NERVOUS SYSTEM, EXPRESSED PREDOMINANTLY  
CC BY SCHWANN CELLS. IN THE CENTRAL NERVOUS SYSTEM, PREFERENTIALLY  
CC EXPRESSED BY NEURONAL CELLS.  
CC - INDUCTION: IN RESPONSE TO SCIATIC NERVE INJURY.  
CC - DOMAIN: THE CYTOPLASMIC DOMAIN INTERACTS WITH ENDOPHILIN I AND  
CC SORTING NEXIN 9 (BY SIMILARITY).  
CC - DOMAIN: DISINTEGRIN DOMAIN BINDS TO INTEGRIN ALPHA5V-BETA3 (BY  
CC SIMILARITY).  
CC - PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
CC SIMILARITY).  
CC - PTM: Phosphorylation increases association with PTKs (By  
CC similarity).  
CC - SIMILARITY: Belongs to peptidase family M12B.  
CC - SIMILARITY: Contains 1 EGF-like domain.  
CC - SIMILARITY: Contains 1 disintegrin domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AJ251198; CAB61762.1; -  
DR HSSP: P17494; 1KST.  
DR MEMOPS: M12.215; -  
DR InterPro: IPR006586; ADAM cysteine.  
DR InterPro: IPR001762; Disintegrin.  
DR InterPro: IPR006209; EGF like.  
DR InterPro: IPR001818; Matrxin.  
DR InterPro: IPR002870; Peg\_M12B\_propep.  
DR InterPro: IPR001590; ReprOlysin.  
DR InterPro: IPR006025; Zn\_MTPeptide.  
DR Pfam: PF00200; disintegrin; 1.  
DR Pfam: PF01562; Peg\_M12B\_propep; 1.  
DR Pfam: PF01421; ReprOlysin; 1.  
DR ProDom: PD000664; Disintegrin; 1.  
DR SMART: SM00608; ACR; 1.  
DR SMART: SM00050; DISIN; 1.  
DR PROSITE: PS50215; ADAM\_MEMPRO; 1.  
DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
DR PROSITE: PS50214; DISINTEGRIN\_2; 1.  
DR PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE: PS01186; EGF\_2; 1; FALSE\_NEG.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE: PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.

KW Hydroxylase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Transmembrane; EGF-like domain; SH3-binding; Phosphorylation.  
 FT SIGNAL 1 17  
 FT PROPEP 18 208  
 FT CHAIN 209 816  
 FT DOMAIN 209 698  
 FT TRANSMEM 699 719  
 FT DOMAIN 720 816  
 FT DOMAIN 209 416  
 FT DOMAIN 423 510  
 FT DOMAIN 511 658  
 FT DOMAIN 659 687  
 FT DOMAIN 700 713  
 FT SITE 768 774  
 FT SITE 803 809  
 FT SITE 180 180  
 FT METAL 350 350  
 FT ACT\_SITE 351 351  
 FT METAL 354 354  
 FT METAL 360 360  
 FT DISULFID 325 411  
 FT DISULFID 482 495  
 FT DISULFID 659 669  
 FT DISULFID 663 673  
 FT DISULFID 677 686  
 FT MOD\_RES 717 717  
 FT MOD\_RES 737 737  
 FT CARBOHYD 57 57  
 FT CARBOHYD 239 239  
 FT CARBOHYD 391 391  
 FT CARBOHYD 394 394  
 FT CARBOHYD 608 608  
 FT CARBOHYD 613 613  
 SQ SEQUENCE 816 AA; 88051 MW; B9D2CE023266FC27 CRC64;

Query Match 52.3%; Score 173.5; DB 1; Length 816;  
 Best Local Similarity 53.7%; Pred. No. 4,4e-11;  
 Matches 29; Conservative 8; Mismatches 14; Indels 3; Gaps 2;

OY 1 CSLSNGAHC-SDGPCNNNTSCLPQPGYECEDANMEDITFCYCTGDSGCCPPNL 53  
 DB 455 COLPFGACASDGGCCCN--CKLQPAQWQRLPTDDCDLPERCLDSSQCPDI 506

RESULT 8  
 AD15\_HUMAN STANDARD; PRT; 814 AA.  
 AC Q13444; Q13493; Q96C78;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE ADAM15 precursor (BC 3-4-24-~) (A disintegrin and metalloprotease domain 15) (Metalloprotease-like, disintegrin-like, and cysteine-rich protein 15) (MDC-15) (Metalloprotease RGD disintegrin protein) (Metargidin).  
 DE ADAM15 OR MDC15.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=96214870; PubMed=8617717;  
 RA Knaetzschmar J., Lum L., Blobel C.P.;  
 RT "Metargidin, a membrane-anchored metalloprotease-disintegrin protein with an RGD integrin binding sequence.";  
 RL J. Biol. Chem. 271:4593-4596(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Umbilical vein;  
 RX MEDLINE=97192141; PubMed=9039960;  
 RA Herren B., Raines E.W., Ross R.;

RT "Expression of a disintegrin-like protein in cultured human vascular  
 RT cells and in vivo.";  
 RL FASEB J. 11:173-180(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Karkkainen I., Huovila A.-P.J.;  
 RT "The characterization of human ADAM15 gene structure and promoter  
 RT region.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,  
 RA Brownstein M.J., Udell T.B., Tohiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP INTERACTION WITH INTEGRIN ALPHA-V-BETA3.  
 RX MEDLINE=98194837; PubMed=9516430;  
 RA Zhang X.P., Kamata T., Yokoyama K., Puzon-McLaughlin W., Takada Y.;  
 RT "Specific interaction of the recombinant disintegrin-like domain of  
 RT MDC-15 (metargidin, ADAM-15) with integrin alphavbeta3.";  
 RL J. Biol. Chem. 273:7345-7350(1998).  
 RN [6]  
 RP PHOSPHORYLATION OF TYR-715 AND TYR-735.  
 RX MEDLINE=21826475; PubMed=11741929;  
 RA Poghosyan Z., Robbins S.M., Houslay M.D., Webster A., Murphy G.,  
 RA Edwards D.R.;  
 RT "Phosphorylation-dependent interactions between ADAM15 cytoplasmic  
 RT domain and Src family protein-tyrosine kinases.";  
 RL J. Biol. Chem. 277:4999-5007(2002).  
 CC -1- FUNCTION: May be involved in cell-surface proteolysis, cell  
 CC adhesion or intracellular protein maturation.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBUNIT: Interacts with ITGA5-ITGB3 (vitronectin receptor),  
 CC PACSIN3 and SNX9. PACSIN3 and SNX9 preferentially bind the  
 CC precursor but not the processed form of ADAM15, suggesting that  
 CC the interaction occurs in a secretory pathway compartment prior to  
 CC the medial Golgi (By similarity). Interacts specifically with Src  
 CC family protein-tyrosine kinases (PTKs).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Ubiquitously EXPRESSED. OVEREXPRESSED IN  
 CC ARTERIOSCLEROTIC LESIONS. CONSTITUTIVELY EXPRESSED IN CULTURED  
 CC ENDOTHELIAL AND SMOOTH MUSCLE.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN INTERACTS WITH ENDOPHLIN I AND  
 CC SORTING NEXTIN 9 (BY SIMILARITY).  
 CC -1- DOMAIN: DESINTEGRIN DOMAIN BINDS TO INTEGRIN ALPHA-V-BETA3.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- PTM: Phosphorylation increases association with PTKs.  
 CC -1- SIMILARITY: Belongs to peptidase family M2B.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -







DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN 1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN 2; 1.  
 DR PROSITE; PS00022; EGF 1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF 2; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.  
 DR Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Transmembrane; EGF-like domain; Alternative splicing; Repeat.  
 KM SIGNAL  
 FT PROPEP 1 27  
 FT CHAIN 28 198  
 FT DOMAIN 199 790  
 FT TRANSSEM 199 687  
 FT DOMAIN 688 708  
 FT DOMAIN 709 790  
 FT DOMAIN 199 393  
 FT DOMAIN 399 485  
 FT DOMAIN 486 632  
 FT DOMAIN 629 663  
 FT DOMAIN 732 776  
 FT REPEAT 732 740  
 FT REPEAT 741 749  
 FT REPEAT 750 758  
 FT REPEAT 759 767  
 FT REPEAT 768 776  
 FT SITE 172 172  
 FT METAL 338 338  
 FT ACT\_SITE 339 339  
 FT METAL 342 342  
 FT METAL 348 348  
 FT DISULFID 313 388  
 FT DISULFID 353 373  
 FT DISULFID 457 470  
 FT DISULFID 633 644  
 FT DISULFID 638 650  
 FT DISULFID 652 661  
 FT CARBOHYD 222 222  
 FT CARBOHYD 372 372  
 FT CARBOHYD 438 438  
 FT CARBOHYD 473 473  
 FT CARBOHYD 625 625  
 FT VARSPLIC 763 771  
 FT FT  
 FT CONFLICT 336 336  
 FT CONFLICT 359 359  
 SQ SEQUENCE 790 AA; 88923 MM; 88EEF8215E4A84F0 CMC64;  
 Query Match 50.3%; Score 167; DB 1; Length 790;  
 Best Local Similarity 50.9%; Pred. No. 2.1e-10;  
 Matches 28; Conservative 9; Mismatches 16; Indels 2; Gaps 1;  
 Oy 1 CSLNSGACSDGPPCCNNTSCLPFGPGYCRDAVNECDITCYCTGSGCCPNMJK 55  
 Db 431 CKLPGANCSTGLCCHD--CRRPBGVYCRQEGNECDLAECYDGNSSCPNDVYK 483  
 RESULT 10  
 AD33\_HUMAN STANDARD; PRT; 813 AA.  
 AC Q9B211;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAM 33 precursor (EC 3.4.24.-) (A disintegrin and metalloprotease  
 domain 33).  
 GN ADAM33.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC Tissue=Testis;

RX MEDLINE=21674006; PubMed=11814695;  
 RA Yoshinaka T., Nishii K., Yamada K., Sawada H., Nishiwaki E., Smith K.,  
 RA Yoshino K., Ishiguro H., Yamashiyama S.;  
 RT "Identification and characterization of novel mouse and human ADAM33s  
 RL with potential metalloprotease activity.";  
 RL Gene 282:227-236 (2002).  
 RP SEQUENCE OF 86-813 FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Scriver G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beate D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharasliho M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McClay K., McMurtry A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sultun J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitteker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871 (2001).  
 CC -1- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9B211-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9B211-2; Sequence=VSP\_005495;  
 CC -1- TISSUE SPECIFICITY: Expressed in all tissues, except liver, with  
 CC high expression in placenta, lung, spleen and veins.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 disintegrin domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB055891; BAB83092.1; .  
 CC EMBL; AL361755; CAC16509.2; .  
 CC HSSP; P1619; 1FVL.  
 CC MEROPS; M12.239; .  
 CC Genew; HGNC:15478; ADAM33.  
 CC MIM; 607114; .  
 CC GO; GO:0016021; C:integral to membrane; NAS.  
 CC GO; GO:0004232; F:metalloendopeptidase activity; NAS.  
 CC GO; GO:0008270; F:zinc ion binding activity; NAS.  
 CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
 CC GO; GO:0042127; P:regulation of cell proliferation; ISS.  
 CC GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.  
 CC InterPro; IPR006586; ADAM\_cysteine.

```
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001818; Matrxin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR006025; Zn_M12B_propep.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MERO; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
KW Transmembrane; EGF-like domain; Alternative splicing.
FT SIGNAL; 1 29
FT PROPEP; 30 203
FT CHAIN; 204 813
FT DOMAIN; 30 701
FT TRANSMEM; 702 722
FT DOMAIN; 723 813
FT DOMAIN; 204 409
FT DOMAIN; 417 502
FT DOMAIN; 503 648
FT DOMAIN; 649 681
FT SITE; 133 133
FT METAL; 345 345
FT ACT_SITE; 346 346
FT METAL; 349 349
FT METAL; 355 355
FT DISULFID; 320 404
FT DISULFID; 360 388
FT DISULFID; 475 488
FT DISULFID; 653 663
FT DISULFID; 657 669
FT DISULFID; 671 680
FT CARBOHYD; 109 109
FT CARBOHYD; 145 145
FT CARBOHYD; 231 231
FT CARBOHYD; 276 276
FT CARBOHYD; 448 448
FT VARSPLIC; 636 661
SQ SEQUENCE 813 AA; 87738 MW; 90713A95668D5569 CRC64;

Query Match 48.2%; Score 160; DB 1; Length 813;
Best Local Similarity 48.1%; Pred. No. 1.le-09;
Matches 26; Conservative 9; Mismatches 17; Indels 2; Gaps 1;

OY 1 CSLNSGAHCSDPCCNNTSCLPFRGYECRDVNECDITETCTGDSGQCPNLIH 54
DB 449 CSLRPGAGCAHGDCC--VRCLLKPAAGALCRQAMGDCDDLPEFCTGTSSHCPDVI 500
```

Search completed: October 21, 2003, 10:06:25  
Job time : 10.6855 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 09:53:36 ; Search time 38.3962 Seconds  
(without alignments)  
369.643 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586

Perfect score: 332  
Sequence: 1 CSLNGAHCHSDGPCNNNTSC.....CDITEYCTGDSGQCPPLHAK 55

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioph:\*  
17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	832	4	075077
2	319	96.1	690	11	08CCJ3
3	319	96.1	829	11	09RI17
4	173.5	52.3	81	11	P97778
5	168.5	50.8	814	4	096C78
6	168	50.6	617	13	090499
7	167	50.3	790	4	08TB27
8	165	49.7	836	6	019057
9	162	48.8	694	5	09G215
10	162	48.8	1182	4	09VXL1
11	160	48.2	812	4	08N0W6
12	158.5	47.7	809	11	08CAB2
13	158.5	47.7	864	11	08C720
14	154.5	46.5	812	6	077779
15	153	46.1	736	6	028482
16	150	45.2	873	13	042595

17	149.5	45.0	212	13	090144	09044 agkistron
18	149.5	45.0	610	13	09Y120	09Y120 agkistron
19	149	44.9	914	13	012960	012960 xenopus lae
20	148	44.6	791	11	060813	060813 mus musculu
21	148	44.6	804	11	060410	060410 cavia porce
22	147	44.3	451	13	09PT49	09PT49 atractaspis
23	146.5	44.1	609	13	09M6M5	09M6M5 agkistron
24	146	44.0	792	6	019061	019061 saginus oe
25	146	44.0	821	6	019060	019060 coturnix oe
26	146	44.0	922	13	08UVF2	08UVF2 coturnix co
27	145.5	43.8	609	13	090282	090282 crocalus at
28	145.5	43.8	610	13	093523	093523 boctrops ja
29	145.5	43.8	610	13	08QG88	08QG88 boctrops ja
30	145	43.7	789	11	P70505	P70505 ractus norv
31	144.5	43.5	587	5	08T319	08T319 drosophila
32	144.5	43.5	959	5	09VPH0	09VPH0 drosophila
33	144	43.4	825	6	046552	046552 papio anubi
34	144	43.4	825	6	028477	028477 macaca fasc
35	144	43.4	838	6	019056	019056 papio anubi
36	144	43.4	905	6	028476	028476 macaca fasc
37	143	43.1	620	13	042138	042138 agkistron
38	143	43.1	702	11	08BMR5	08BMR5 mus musculu
39	143	43.1	714	11	08K4K0	08K4K0 mus musculu
40	141	42.5	797	11	08R465	08R465 mus musculu
41	141	42.5	797	11	08R5G5	08R5G5 mus musculu
42	140.5	42.3	190	13	093518	093518 agkistron
43	140	42.2	697	11	08BMR4	08BMR4 mus musculu
44	139.5	42.0	761	11	08CDV3	08CDV3 mus musculu
45	139	41.9	845	11	061072	061072 mus musculu

## ALIGNMENTS

RESULT 1  
ID 075077 PRELIMINARY; PRT; 832 AA.  
AC 075077;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE MDC3 (ADAM22 protein).  
GN ADAM22.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98359734; PubMed=9693107;  
RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.;  
RT "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2  
RT and MDC3: novel human cellular disintegrins highly expressed in the  
RT brain.";  
RL Biochem. J. 334:93-98(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Cal S., Lopez-Otin C.;  
RT "Identification and characterization of ADAM 22, a novel  
RT metalloproteinase/disintegrin protein.";  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB009672; BAA32351.1; -;  
DR HSSP; AJ005580; CAC20565.1; -;  
DR HSSP; P18619; IFVL.  
DR MEROPS; M2.979; -;  
DR Genew; HGNC:202; ADAM23.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR00162; Disintegrin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002870; Pep\_M12B\_propep.

```
DR InterPro; IPR001590; Reprolysin.
DR Pfam; PF00200; disintegrin. 1.
DR Pfam; PF01562; Pep M12B_propep. 1.
DR Pfam; PF01421; Reprolysin. 1.
DR ProDom; PD000664; Disintegrin. 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
SQ SEQUENCE 832 AA; 91925 MW; 78419670E1C24EF CRC64;

Query Match
Best Local Similarity 100.0%; Score 332; DB 4; Length 832;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 55
DB 532 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 586

RESULT 2
Q8CC33 PRELIMINARY; PRT; 690 AA.
AC Q8CC33;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE A disintegrin and metalloprotease domain 23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=dienecephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RL EMBL; AK034022; BAC28550.1;
SQ SEQUENCE 690 AA; 77228 MW; B61699141A6CA6A6 CRC64;

Query Match
Best Local Similarity 96.4%; Score 319; DB 11; Length 690;
Matches 53; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 55
DB 529 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 583

RESULT 3
Q9R1V7 PRELIMINARY; PRT; 829 AA.
AC Q9R1V7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ADAM23.
GN ADAM23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99365303; PubMed=10433968;
RX Segane K., Yamazaki K., Mizui Y., Tanaka I.;
```

```
RT "Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23.";
RL Gene 236:79-86(1999).
DR EMBL; AB009673; BAAB3381.1; -
DR HSSP; P18619; 1FVL.
DR MEROPS; M12.979; -.
DR MGD; MGI:1345162; Adam23.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR Pfam; PF00200; disintegrin. 1.
DR Pfam; PF01562; Pep M12B_propep. 1.
DR Pfam; PF01421; Reprolysin. 1.
DR ProDom; PD000664; Disintegrin. 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
SQ SEQUENCE 829 AA; 91547 MW; FE6BCD69DD50F53A CRC64;

Query Match
Best Local Similarity 96.1%; Score 319; DB 11; Length 829;
Matches 53; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 55
DB 529 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 583

RESULT 4
P97778 PRELIMINARY; PRT; 81 AA.
AC P97778;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TMDc VI (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague Dawley; TISSUE=Testis;
RX MEDLINE=97436783; PubMed=9291465;
RA Frayne J., Jury J.A., Barker H.L., Hall L.;
RT "Sequence analysis, tissue distribution, and expression in prepubertal
RT and adult rat testis."
RL Mol. Reprod. Dev. 48:159-167(1997).
RL EMBL; Y11492; CAA72278.1; -.
DR HSSP; P17494; 1KST.
DR MEROPS; M12.215; -.
DR InterPro; IPR001762; Disintegrin.
DR Pfam; PF00200; disintegrin. 1.
DR ProDom; PD000664; Disintegrin. 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 8622 MW; 104013B27E1096CB CRC64;
```

```
Query Match
Best Local Similarity 52.3%; Score 173.5; DB 11; Length 81;
Matches 29; Conservative 8; Mismatches 14; Indels 3; Gaps 2;

QY 1 CSLSNAGHCSGDPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 53
DB 16 COLRRGACASDGPCCNNTSCLFQPRGYECRDVAVNECDITEYCTGDSGQCPPLHK 67
```

```

RESULT 5
ID 096C78 PRELIMINARY; PRT; 814 AA.
AC 096C78;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE A disintegrin and metalloproteinase domain 15 (metargidin)
DE Metalloprotease disintegrin.
GN ADAM15.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strauberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Karkkainen I., Huovila A.-P.J.;
RT "The characterization of human ADAM15 gene structure and promoter
region."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014566; AAH14566.1; -.
DR EMBL: AF314227; AAM44189.1; -.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR InterPro: IPR006025; Zn_MTPpeptase.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Repolysin; 1.
DR ProDom: PD000664; Disintegrin; 1.
DR PROSITE: PS50215; ADAM_MERO; 1.
DR PROSITE: PS50214; DISINTEGRIN_2; 1.
DR PROSITE: PS00186; EGF_2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Integrin; Protease.
SQ SEQUENCE 814 AA; 8717 MW; 683A8368AD3096B CRC64;

Query Match 50.8%; Score 168.5; DB 4; Length 814;
Best Local Similarity 51.9%; Pred. No. 2.2e-14;
Matches 28; Conservative 8; Mismatches 15; Indels 3; Gaps 2;

OY 1 CSLSNGAHC-SDGFCNNNTSCLFQPRGYECRDVAIVEDITEYCTGDSGCCPNTL 53
Db 453 CQLRPGAGCAGSDGFCQCN-CQLRPGAGCQCPTRGDDCLPFCFCDSSDCCPDV 504

RESULT 6
ID 090499 PRELIMINARY; PRT; 617 AA.
AC 090499;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Metalloprotease.
GN ECHI.
OS Echin pyramidium (carpet viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Echis.
OX NCBI_Taxid=8700;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=95010025; PubMed=7925363;
RA Paine M.V.I., Moura-Da-Silva A.M., Theakston R.D.G., Crampton J.M.;
"Cloning of metalloprotease genes in the carpet viper (Echis pyramidium

```

```

RT (leakey): Further members of the metalloprotease/disintegrin gene
family."
RL Eur. J. Biochem. 224:483-486(1994).
DR EMBL: X78970; CA555565.1; -.
DR HSP: P18619; LFVL.
DR InterPro: IPR006586; ADAM_cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR InterPro: IPR006025; Zn_MTPpeptase.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Repolysin; 1.
DR PRINTS: PR00289; DISINTEGRIN.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00608; ACR; 1.
DR SMART: SM00050; DISIN; 1.
DR PROSITE: PS50215; ADAM_MERO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; 1.
DR PROSITE: PS50214; DISINTEGRIN_2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Protease.
FT CHAIN 193 617 METALLOPROTEASE.
SQ SEQUENCE 617 AA; 69310 MW; 83DC3DA5F4F3AFBA CRC64;

Query Match 50.6%; Score 168; DB 13; Length 617;
Best Local Similarity 52.0%; Pred. No. 1.9e-14;
Matches 26; Conservative 9; Mismatches 13; Indels 2; Gaps 1;

OY 1 CSLSNGAHCSDGFCNNNTSCLFQPRGYECRDVAIVEDITEYCTGDSGCCP 50
Db 436 CKLTPGSCADGECQCN-CQFRFPARTECRKIDCDVPEYCTGSGSECP 483

RESULT 7
ID 08TB27 PRELIMINARY; PRT; 790 AA.
AC 08TB27;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE A disintegrin and metalloproteinase domain 30.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC028372; AAH28372.1; -.
DR InterPro: IPR006586; ADAM_cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Repolysin.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Repolysin; 1.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00608; ACR; 1.
DR SMART: SM00050; DISIN; 1.
DR PROSITE: PS50215; ADAM_MERO; 1.
DR PROSITE: PS50214; DISINTEGRIN_2; 1.
DR PROSITE: PS00186; EGF_2; 1.
DR Integrin.
SQ SEQUENCE 790 AA; 88940 MW; 42EC8A5F66ECCDA3 CRC64;

Query Match 50.3%; Score 167; DB 4; Length 790;
Best Local Similarity 50.9%; Pred. No. 3.4e-14;
Matches 28; Conservative 9; Mismatches 16; Indels 2; Gaps 1;

```

```
Oy 1 CSLSNGAHCSGDPCCNNNTSCLEFQPRGYECRDVAVNECDITEYCTGDSGCCPPNLRK 55
DB 431 CKLQAGANCSTGLCHD--CRFRPSGYVCROGNCNCDLAEYCDGSSSCPNVYK 483

RESULT 8
ID 019057 PRELIMINARY; PRT; 836 AA.
AC 019057;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Fertilin alpha protein.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.
OC NCB1_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Blood;
RX MEDLINE=98376167; PubMed=9712322;
RA Juy J.A., Frayne J., Hall L.;
RT "Sequence analysis of a variety of primate fertilin alpha genes:
RT Evidence for non-functional genes in the gorilla and man.";
RL Mol. Reprod. Dev. 51:92-97(1998).
DR EMBL; Y15491; CAA75659.1; -.
DR HSSP; P18619; 1FVL.
DR MEROPS; M12_201; -.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR InterPro; IPR006025; Zn_MTpeptidase.
DR Pfam; PF00200; disintegrin_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MERPO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 836 AA; 93135 MW; 659CDE6A5991BA12 CRC64;

Query Match 49.7%; Score 165; DB 6; Length 836;
Best Local Similarity 50.9%; Pred. No. 6,8e-14;
Matches 28; Conservative 5; Mismatches 20; Indels 2; Gaps 1;

Oy 1 CSLSNGAHCSGDPCCNNNTSCLEFQPRGYECRDVAVNECDITEYCTGDSGCCPPNLRK 55
DB 473 CKLKAACSCDGPCH--KCKFQKRGKYPCCPSRSSCDLPFCNGTSALCPDNRHK 525

RESULT 9
ID 09G215 PRELIMINARY; PRT; 694 AA.
AC 09G215;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE MIND-WELD (Fragment).
GN MMD OR CG9163 OR CG15603 OR CG15604.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Harper C.H., Chase B.A.;
RT "The Drosophila mind-weld gene encodes a neuronally expressed ADAM
RT protein.";
RA Submitted (Apr-2000) to the EMBL/Genbank/DDBJ databases.
```

```
DR EMBL; AF252287; AAF98331.1; -.
DR HSSP; P18619; 1FVL.
DR FlyBase; FBgn0041109; mmd.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Repolysin.
DR Pfam; PF00200; disintegrin_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS50215; ADAM_MERPO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW EGF-like domain.
FT NON-TER 1
SQ SEQUENCE 694 AA; 76317 MW; 37736F7B064A72E9 CRC64;

Query Match 48.8%; Score 162; DB 5; Length 694;
Best Local Similarity 47.3%; Pred. No. 1.4e-13;
Matches 26; Conservative 10; Mismatches 17; Indels 2; Gaps 1;

Oy 1 CSLSNGAHCSGDPCCNNNTSCLEFQPRGYECRDVAVNECDITEYCTGDSGCCPPNLRK 55
DB 369 CKLSEAGCAGACCD--QCRLRPKDYICRNSNNECDLPXCDEIDIGCPEDVYK 421

RESULT 10
ID 09VXL1 PRELIMINARY; PRT; 1182 AA.
AC 09VXL1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CG9163 protein.
GN MMD OR CG9163 OR CG15603 OR CG15604.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
RA Ballaw R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beas P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotiler P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hooten D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
```

RA Laeko P., lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard R., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).

RP [12] SEQUENCE FROM N.A.

RA Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorselt V., Doup L.E., Doyle C., Dresnek D., Fafan D.,  
 RA Ferrera S., Frise E., Galle R.F., Gary N.S., George H.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Paclet J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [13] SEQUENCE FROM N.A.

RA Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hrdecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Ceiniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [14] SEQUENCE FROM N.A.

RP Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [15] SEQUENCE FROM N.A.

RA FlyBase; submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AE003500; AAF48548.2; -.  
 DR HSSP; P18619; 1FVL.  
 DR FlyBase; FBgn0041109; mmd.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR001590; Reptolysin.  
 DR Pfam; PF00200; disintegrin; 1.  
 DR Pfam; PF01421; Reptolysin; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR PRODOM; PD000664; Disintegrin; 1.  
 DR SMART; SM00050; DISIN; 1.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 SO SEQUENCE 1182 AA; 128515 MW; F58BA6E349C2ECA9 CRC64;

Query Match 48.8%; Score 162; DB 5; Length 1182;  
 Best Local Similarity 47.3%; Pred. No. 2,5e-13;  
 Matches 26; Conservative 10; Mismatches 17; Indels 2; Gaps 1;

Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDVAWNECDITEYCTGDSGQCPPLHK 55  
 Db 231 CKLSEAKQASGACD--QCRLAPKQYICRDSNNECDLPRYCQGEICQCCSDVFK 283

Search completed: October 21, 2003, 10:10:36  
 Job time : 40.3962 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:12:31 ; Search time 3432.65 Seconds  
(without alignment)  
1239.449 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599  
Perfect score: 630  
Sequence: 1 PTKLFEPTCEGNGYVEAGEE.....CPMLHKQDQACNONGRC 104

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh  
-Q/cgnt2\_1/USPRO.spool/US09634252/runat\_21102003\_090323\_3370/app\_query.fasta\_1.462  
-DB=GenEmb1 -QPM=faetap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOCPCT=0 -LOOPEXT=0  
-UNITS=b1b -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09634252 @CGN 1 1 3608 @runat\_21102003\_090323\_3370 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLIST  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

1: GenEmb1:\*  
2: gb\_ba:\*  
3: gb\_neg:\*  
4: gb\_in:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_pln:\*  
35: em\_hlg\_rtd:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vtr:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	2499	6 BD130423	BD130423 DNs and
2	630	100.0	3054	6 E26532	E26532 Novel prote
3	630	100.0	3054	9 AB009672	AB009672 Homo sapi
4	630	100.0	3079	9 HSAS580	AJ005580 Homo sapi
5	630	100.0	4043	6 AX299710	AX299710 Sequence
6	611	97.0	2891	10 AB009673	AB009673 Mus muscu
7	586	93.0	1668	6 AX235014	AX235014 Sequence
8	404.5	64.2	1464	6 E10490	E10490 CDNA encodi
9	404.5	64.2	1464	6 E12845	E12845 Sequence 5
10	404.5	64.2	1464	6 E13482	E13482 Sequence 5
11	404.5	64.2	2908	6 HDMDC	D17390 Homo sapien
12	404.5	64.2	2913	6 E10492	E10492 CDNA encodi
13	404.5	64.2	2913	6 E12847	E12847 Sequence 7
14	404.5	64.2	2913	6 E13484	E13484 Sequence 7
15	404.5	64.2	2923	6 E10491	E10491 CDNA encodi
16	404.5	64.2	2923	6 E12846	E12846 Sequence 6
17	404.5	64.2	2923	6 E13483	E13483 Sequence 6
18	404.5	64.2	3168	9 AB009675	AB009675 Homo sapi
19	404.5	64.2	3168	6 E10493	E10493 CDNA encodi
20	404.5	64.2	3163	6 E12848	E12848 Sequence 8
21	404.5	64.2	3183	6 E13485	E13485 Sequence 8
22	404.5	64.2	3229	10 AB009676	AB009676 Mus muscu
23	375	59.5	1588	5 AF032384	AF032384 Xenopus 1
24	364.5	57.9	2773	10 AB009674	AB009674 Mus muscu
25	363.5	57.7	2604	6 BD130424	BD130424 DNs and
26	363.5	57.7	2697	6 E26531	E26531 Novel prote
27	363.5	57.7	2786	9 AB009671	AB009671 Homo sapi
28	363.5	57.7	2796	9 AF155381	AF155381 Homo sapi
29	363.5	57.7	2805	6 E26530	E26530 Novel prote
30	363.5	57.7	2858	9 AF155382	AF155382 Homo sapi
31	363.5	57.7	3259	9 AF158637	AF158637 Homo sapi
32	363.5	57.7	3295	9 AF073291	AF073291 Homo sapi
33	357.5	56.7	3447	5 AF032383	AF032383 Xenopus 1
34	338.5	53.7	1674	6 AX235012	AX235012 Sequence
35	292	46.3	2541	9 AB055891	AB055891 Homo sapi
36	292	46.3	2642	6 AX686216	AX686216 Sequence
37	292	46.3	2649	6 AX319859	AX319859 Sequence
38	292	46.3	2705	6 AX686214	AX686214 Sequence
39	292	46.3	2762	6 AX686212	AX686212 Sequence
40	292	46.3	3033	6 HSM601104	AL117415 Homo sapi
41	292	46.3	3431	6 AR219232	AR219232 Sequence
42	292	46.3	3431	6 AX082151	AX082151 Sequence
43	292	46.3	3468	6 AR219233	AR219233 Sequence
44	292	46.3	3468	6 AX082153	AX082153 Sequence
45	292	46.3	3509	6 AX326807	AX326807 Sequence

RESULT 1

#### ALIGNMENTS

BD130423  
LOCUS BD130423 2499 bp DNA linear PAT 18-SEP-2002  
DEFINITION DNAe and polypeptides of metalloprotease disintegrins SVPH3-13 and SVPH3-17.  
ACCESSION BD130423  
VERSION BD130423.1 GI:23225368  
KEYWORDS JP 2002503472-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 2499)  
AUTHORS Cerretti,D.P.  
TITLE DNAs and polypeptides of metalloprotease disintegrins SVPH3-13 and SVPH3-17  
JOURNAL Patent: JP 2002503472-A 2 05-FEB-2002;  
IMMUNEX CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002503472-A/2  
PD 05-FEB-2002  
PR 11-FEB-1999 JP 200531569  
PR 11-FEB-1998 US 60/074310  
PI DOUGLAS PAT CERRETTI  
PC C12N15/09,C07K16/40,C12N1/19,C12N1/21,C12N5/10,C12N9/64 PC  
,C12Q1/37,G01N33/68,  
CC C12N15/00,C12N5/00  
CC DNAs and polypeptides of metalloprotease disintegrins CC  
SVPH3-13 and SVPH3-17  
FH Key  
FT source  
FT Location/Qualifiers  
1. 2499  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 677 a 593 c 645 g 584 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3 41e-56 Length: 2499  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x BD130423 (1-2499)

QY 1 ProthrlYsLeuPhgGluProThrgIuCySGIyAsnGlyTYrVAIGluAglYgluGlu 20  
DB 1486 CCAACAAAGCTATTGAGCCCAAGGAAATGGAATGATACGTGAAGCTGGGAGAG 1545  
QY 21 CysAaPcYsgIyPheHisValaGluCySTyrgIyLeuCyCyAluYsCySerIeuSer 40  
DB 1546 TGTGATTGTGGTTTCATGTCGAATGCTATGATTAATGCTGAAGAAATGTTCCCTCC 1605  
QY 41 AsnGIyAlaHisCySerAspGlyProCyCyAsnAsnThSerCySeuPheGlnPro 60  
DB 1606 AACGGGGCTCACTGACGACGAGCGGCCCTGCTGAACAAATACCTCATGCTTTTTCAGCCA 1665  
QY 61 ArgGIyTYrGluCyArGAspAlaValaAsnGluCyAspIleThGluTYrCySThrgIy 80  
DB 1666 CGAGGGTATGAATGCGGGGATGCTGTGAACGAGTGAATTAATTAATTTGTAATGTA 1725  
QY 81 AspSerGIyGlnCySProProAsnLeuHisIySgIAspGIyTYrAlaCyAsnGlnAsn 100  
DB 1726 GACTCTGTGCTAGTGCACCAATATCTTCAATAGCAAGAGATATGATGCAATCAAAAT 1785  
QY 101 GlnGIyArGcyS 104  
DB 1786 CAGGGCCGCTGC 1797

RESULT 2  
LOCUS E26532 3054 bp DNA linear PAT 18-JUN-2001  
DEFINITION Novel protein belonging to MDC gene family and DNA encoding the same.  
ACCESSION E26532  
VERSION E26532.1 GI:13026199  
KEYWORDS JP 1999155574-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 3054)  
AUTHORS Koji,S.  
TITLE Novel protein belonging to MDC gene family and DNA encoding the same  
JOURNAL Patent: JP 1999155574-A 3 15-JUN-1999;  
EISAI CO LTD  
COMMENT OS Homo sapiens (human)  
PN JP 1999155574-A/3  
PD 15-JUN-1999  
PR 01-DEC-1997 JP 1997330020  
PI KOJI SONE  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/21,C12P21/02,C12Q1/68// PC  
(C12N15/09,C12R1/91),(C12N1/21,C12R1/19),(C12P21/02,C12R1/19), PC  
C12N15/00,  
CC (C12N15/00,C12R1/91)  
CC Strandedness: Double;  
CC Topology: linear;  
FH Key  
FT CDS  
FT Location/Qualifiers  
1. 3054  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 815 a 754 c 803 g 682 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4 21e-56 Length: 3054  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x E26532 (1-3054)

QY 1 ProthrlYsLeuPhgGluProThrgIuCySGIyAsnGlyTYrVAIGluAglYgluGlu 20  
DB 1709 CCAACAAAGCTATTGAGCCCAAGGAAATGGAATGATACGTGAAGCTGGGAGAG 1768  
QY 21 CysAaPcYsgIyPheHisValaGluCySTyrgIyLeuCyCyAluYsCySerIeuSer 40  
DB 1769 TGTGATTGTGGTTTCATGTCGAATGCTATGATTAATGCTGAAGAAATGTTCCCTCC 1828  
QY 41 AsnGIyAlaHisCySerAspGlyProCyCyAsnAsnThSerCySeuPheGlnPro 60  
DB 1829 AACGGGGCTCACTGACGACGAGCGGCCCTGCTGAACAAATACCTCATGCTTTTTCAGCCA 1888  
QY 61 ArgGIyTYrGluCyArGAspAlaValaAsnGluCyAspIleThGluTYrCySThrgIy 80  
DB 1889 CGAGGGTATGAATGCGGGGATGCTGTGAACGAGTGAATTAATTAATTTGTAATGTA 1948  
QY 81 AspSerGIyGlnCySProProAsnLeuHisIySgIAspGIyTYrAlaCyAsnGlnAsn 100  
DB 1949 GACTCTGTGCTAGTGCACCAATATCTTCAATAGCAAGAGATATGATGCAATCAAAAT 2008  
QY 101 GlnGIyArGcyS 104  
DB 2009 CAGGGCCGCTGC 2020

```
RESULT 3
AB009672      3054 bp      mRNA      linear      PRI 15-AUG-1998
DEFINITION    Homo sapiens mRNA for MDC3, complete cds.
ACCESSION     AB009672
VERSION       AB009672.1 GI:3419877
KEYWORDS      MDC3.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         1 (bases)
Sagane, K., Ohya, Y., Hasegawa, Y. and Tanaka, I.
MDC3 and MDC3: novel human cellular disintegrins highly expressed
in the brain
Biochem. J. 334 (Pt 1), 93-98 (1998)
JOURNAL       Biochem. J. 334 (Pt 1), 93-98 (1998)
MEDLINE       98359734
PUBMED        9693107
REFERENCE     2 (bases 1 to 3054)
AUTHORS       Sagane, K.
TITLE         Direct Submission
Submitted (11-DEC-1997) Koji Sagane, Eisai Co., Ltd., Tsukuba
Research Laboratories, Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635,
Japan (E-mail: k1-sagane@eisai.co.jp, Tel: +81-298-47-5813,
Fax: +81-298-47-5367)
FEATURES
    source
        1..3054
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="brain"
            224..2722
            /codon_start=1
            /product="MDC3"
            /protein_id="BAA32351.1"
            /db_xref="GI:3419877"
            /translation="MKPPSSSRQPPPLACSLAGACGPGRPGASVPASAPARTPPC
            RLLVLVLPLPLAASRPRAWGAAPSAHPMNTAKNLCVLADEPNTIQNSSNIS
            YSNAMOKETLPSRLIYYINQDSSEPYHVLDTKARHQKRNKVAHLQAQSFQEARGS
            KFTLIDLINNGLSDDYVEIHYENGRPOYSKGEHGYOHSIRGVDSKVALSTCNGL
            HGFEDDTFVYMIPELIVHDEKSTGRPHIIOKTLGQYSKQKNTLMERGQDWPLS
            ELQMLKRRKRAVNPISRGIFPEEMKYLEIMIVNDKTYKHSRSHAHNTNFAKSVNLVD
            SIYKEQNTNRVVLVAETWTEKDQIDITTPVQMLHPSKYRORIKQHDAYHLISRV
            TTHYKRSLSYFGVCSRTGVCNMEYGLPMVAOVLSQSLQNLGIOMEPSSRRKXC
            DCTESWGCGIMETGVSHSRKSKSILETRDLQNGACLPNRPKLFEPTECGNG
            YVAGEECDCGFHVECYGLCCCKKCSLSNGAHSDBGCCNNTSCLPQPRGECDAVNE
            CDITEYCTGDSGQCPNLHKQDGYACNONGRCYNBCKTRNDQCYIWTGXAAGSDK
            FCYEKLTGTEGKNGCKDGRWIOCSKHDFGFLICTNLTRAPRIQLOGEIIPTS
            FYHQRVINDSGAHVLLDDPTDYGVEDTGPCSPMWCIDRCLQIOALNMSCPLDS
            KGVCSGHGVCSEATCTCDFTWAGTDCSIRDPVRNLHPKKBGPGPSATNLISIS
            AGAILVAIVLGGTGMGFKNVKRRRDPDTPQGP1"
BASE COUNT   815 a      754 c      803 g      682 t
ORIGIN
Alignment Scores:
    Pred. No.:      4,21e-56      length:      3054
    Score:           630.00      matches:      104
    Percent Similarity: 100.00%      conservative: 0
    Best Local Similarity: 100.00%      mismatches: 0
    Query Match:     100.00%      indels:      0
    DB:              9      gaps:      0
US-09-634-252a-4_COPY_496_599 (1-104) x AB009672 (1-3054)
QY      1 ProThrlYsleuPhegiuProthngluCyseGlyAsnGlyTyValajsluajlgluGlu 20
Db      1709 CCAACAAGCATATTGAGCCACGAGATGTGAAATGATAACGTAAGCTGGGAGAG 1768
QY      21 CyAAspCyseGlyPhehlsvalgluCyseTyGlyleuCyseGlyValylsCyseSerleuSer 40
Db      1769 TGTGATTGTGTTCATGTGGAATGCTATGAGATTATGCTGTAAAGAAATGTTCCCTCTCC 1828
```

```
QY      41 AenglyAlahisCyseSeraspGlyProCyseCyseAsnAsnThrSexCyseuPhegiuPro 60
Db      1829 AACGGGGCTCACTGACGACGCGGCTCTCTTAACAATACTCAATGCTTTTTCAGCCA 1888
QY      61 ArgGlyTYrGluCyseArgaspAlaValaAengluCyseAspIleThrGluTyrcystrGly 80
Db      1889 CGAGGGTATGAATGCGCGGATCTGTGAACGAGTGTGATTAATGATAATATTGCTGGA 1948
QY      81 AspserGlyGluCyseProProbanleuHslYsgInaspGlyTyValaCyseAengluAan 100
Db      1949 GACTCTGGTCAGTGGCCACCAAACTTCTATAGCAAGACGAGATATGACATCAATAAT 2008
QY      101 GInGlyArgCys 104
Db      2009 CAGGCGCGCTGC 2020
RESULT 4
HSA5580
LOCUS         HSA5580      3079 bp      mRNA      linear      PRI 18-JAN-2001
DEFINITION    Homo sapiens mRNA for adam23 protein.
ACCESSION     AJ005580
VERSION       AJ005580.1 GI:12053562
KEYWORDS      adam23 gene; Adam23 protein; disintegrin; metalloprotease.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         Cal, S. and Lopez-Otin, C.
Identification and characterization of ADAM 23, a novel
metalloprotease/disintegrin protein
Unpublished
JOURNAL       2 (bases 1 to 3078)
AUTHORS       Lopez-Otin, C.
TITLE         Direct Submission
Submitted (15-APR-1998) Lopez-Otin C., Bioquimica y Biologia
Molecular, Universidad de Oviedo, C/ Julian Claveria 6, Oviedo.
Asturias, 33006, SPAIN
REMARK        revised by author 23-APR-1998, revised by author 26-OCT-1998
FEATURES
    source
        1..3079
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="brain"
            1..3079
            /gene="adam23"
            <1..45
            /gene="adam23"
            53..2551
            /gene="adam23"
            /function="disintegrin"
            /codon_start=1
            /product="Adam23 protein"
            /protein_id="CAC20565.1"
            /db_xref="GI:12053563"
            /translation="MKPPSSSRQPPPLACSLAGACGPGRPGASVPASAPARTPPC
            RLLVLVLPLPLAASRPRAWGAAPSAHPMNTAKNLCVLADEPNTIQNSSNIS
            YSNAMOKETLPSRLIYYINQDSSEPYHVLDTKARHQKRNKVAHLQAQSFQEARGS
            KFTLIDLINNGLSDDYVEIHYENGRPOYSKGEHGYOHSIRGVDSKVALSTCNGL
            HGFEDDTFVYMIPELIVHDEKSTGRPHIIOKTLGQYSKQKNTLMERGQDWPLS
            ELQMLKRRKRAVNPISRGIFPEEMKYLEIMIVNDKTYKHSRSHAHNTNFAKSVNLVD
            SIYKEQNTNRVVLVAETWTEKDQIDITTPVQMLHPSKYRORIKQHDAYHLISRV
            TTHYKRSLSYFGVCSRTGVCNMEYGLPMVAOVLSQSLQNLGIOMEPSSRRKXC
            DCTESWGCGIMETGVSHSRKSKSILETRDLQNGACLPNRPKLFEPTECGNG
            YVAGEECDCGFHVECYGLCCCKKCSLSNGAHSDBGCCNNTSCLPQPRGECDAVNE
            CDITEYCTGDSGQCPNLHKQDGYACNONGRCYNBCKTRNDQCYIWTGXAAGSDK
            FCYEKLTGTEGKNGCKDGRWIOCSKHDFGFLICTNLTRAPRIQLOGEIIPTS
            FYHQRVINDSGAHVLLDDPTDYGVEDTGPCSPMWCIDRCLQIOALNMSCPLDS
            KGVCSGHGVCSEATCTCDFTWAGTDCSIRDPVRNLHPKKBGPGPSATNLISIS
            AGAILVAIVLGGTGMGFKNVKRRRDPDTPQGP1"
5' UTR
CDS
3' UTR
```

BASE COUNT 865 a 714 c 769 g 731 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,256-56 Length: 3079  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x HSA5580 (1-3079)  
QY 1 Prothrllysleuphegluprothrglucysgllyasnglytyrvalgluaaglygluglu 20  
DB 1538 CCAACAAAGCTATTGGAGCCACCGAATGTGGAATGATACGTGGAAGCTGGAGAG 1597  
QY 21 Cysaapcygsllypnehisvalglucystyrglyleucysylslyscysserleuser 40  
DB 1598 TGTGATTTGTGTTTCATGTGGAATGCTATGATTTATGCTTAAGAAATGTTCCCTCTCC 1657  
QY 41 AsnglylAhlscysserAspglyProCysCysasnaenthrserCysleupheglupro 60  
DB 1658 AACGGGCTCACTGAGGAGCGAGCGGCGCTGCTGTAACAATACCTCATGTTTTTACGCCA 1717  
QY 61 ArgglytyrglucysargaspalaavalasnglucysaspliethrglutyrCysThrgly 80  
DB 1718 CGAGGATATGATGCGCGGATGCTGTGAACGAGTGTGATATTAAGATATTGTTGCTGA 1777  
QY 81 AspsrglyglucysproproasnluehlslysglnaspglytyralaCysasnglnasn 100  
DB 1778 GACTCTGGTCACTGCCACCAAAATCTTCAATAGCAAGCGATATGCAATCAAAAT 1837  
QY 101 GlnglyArgCys 104  
DB 1838 CAGGGCCGCTGC 1849  
RESULT 5 4043 bp DNA linear PAT 26-NOV-2001  
LOCUS AX299710  
DEFINITION Sequence 2 from Patent WO0174857.  
ACCESSION AX299710.1 GI:17129252  
VERSION AX299710.1 GI:17129252  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS Lopez-Otin,C., Freilij,J.M., Bianchi,A.B., Miguel,S.C., Garcia,J.M.  
and Trall,P.  
TITLE Methods and compositions for modulating integrin-mediated cell-cell  
interactions  
JOURNAL Patent: WO 0174857-A 2 11-OCT-2001;  
Bristol-Myers Squibb Co. (US)  
FEATURES  
source 1..4043  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 1110 a 946 c 1002 g 985 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,666-56 Length: 4043  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x AX299710 (1-4043)

QY 1 Prothrllysleuphegluprothrglucysgllyasnglytyrvalgluaaglygluglu 20  
DB 2502 CCAACAAAGCTATTGGAGCCACCGAATGTGGAATGATACGTGGAAGCTGGAGAG 2561  
QY 21 Cysaapcygsllypnehisvalglucystyrglyleucysylslyscysserleuser 40  
DB 2562 TGTGATTTGTGTTTCATGTGGAATGCTATGATTTATGCTTAAGAAATGTTCCCTCTCC 2621  
QY 41 AsnglylAhlscysserAspglyProCysCysasnaenthrserCysleupheglupro 60  
DB 2622 AACGGGCTCACTGAGGAGCGGCGCTGCTGTAACAATACCTCATGTTTTTACGCCA 2681  
QY 61 ArgglytyrglucysargaspalaavalasnglucysaspliethrglutyrCysThrgly 80  
DB 2682 CGAGGATATGATGCGCGGATGCTGTGAACGAGTGTGATATTAAGATATTGTTGCTGA 2741  
QY 81 AspsrglyglucysproproasnluehlslysglnaspglytyralaCysasnglnasn 100  
DB 2742 GACTCTGGTCACTGCCACCAAAATCTTCAATAGCAAGCGATATGCAATCAAAAT 2801  
QY 101 GlnglyArgCys 104  
DB 2802 CAGGGCCGCTGC 2813  
RESULT 6 2891 bp mRNA linear ROD 17-AUG-1999  
LOCUS AB009673  
DEFINITION Mus musculus mRNA for ADAM23, complete cds.  
ACCESSION AB009673.1 GI:5736618  
VERSION AB009673.1 GI:5736618  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS Sagane,K., Yamazaki,K., Mizui,Y. and Tanaka,I.  
TITLE Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23  
JOURNAL Gene 236 (1), 79-86 (1999)  
MEDLINE 99365303  
PIRME 10433368  
REFERENCE 2 (bases 1 to 2891)  
AUTHORS Sagane,K.  
TITLE Direct Submission  
JOURNAL Submitted (11-DEC-1997) Koji Sagane, Eisai Co., Ltd., Tsukuba  
Research Laboratories, Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635,  
Japan (E-mail:k-sagane@hsc.eisai.co.jp, Tel:+81-298-47-5813,  
Fax:+81-298-47-5367)  
COMMENT Sequence updated (07-Aug-1999).  
FEATURES  
source location/Qualifiers  
1..2891  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/issue\_type="brain"  
80..2569  
/codon\_start=1  
/product="ADAM23"  
/protein\_id="BA83381.1"  
/db\_xref="GI:5736618"  
/translation="MKPPGISIRPRLTGGSLPGASCGPGRCPAGVPAPARPCRLLL  
VLLLPALATSSRRARGAAPSAFHNWETAETKTLGLADEDNLTLOONSSRRNTSYS  
AVOKEITLPSRLYINODESPYHVLDTKARHQQKNKVAHLAQSFOAEAPSGKFI  
LDLTNNGLLSSDVEIHYEDGKMYSGKEHCYHNSINGVDSRVALSICNGLHGM  
FEDTFTVYMEPLTDEKSTGRPHITIKTLAQYISKOMKNSITDSDOWPLPELQ  
WLRRKRVNPSRGVFEEMKYLELMTVNDHKTYKRSASHATNNFAKSVNVLVDSLY  
KEQNTKRVVLAVENTEKDIDITIPVQMLHDFSKYRORIKOHADAVHLISRFTY  
YKRSLSYFGVCSIRIGVNEGYGLPMAVOVLOSIAONLGIOWEPSSRKPCECI  
ESMGCIIMEETGVSHSRFSKSTILEYRDLORGGACLFNRPTKLFEPNCGGYVE  
AGEBCDCGFHYECYGVCKKCSUSNGAHCSDGPCNNSTCLFQSRGTGECRDVAVNSCI  
TEYCTGSGQCPNLHRKODGYSNONGRCYNGCKTRDNOCYITWTKAAGSDKFCY  
EKLTBETGERKNGCKDGRWIPCSKHDFGCFLLCTNLTRAPRIGLOGLGELIIPTSFYH

polyA\_site /note="10 a nucleotide"  
BASE COUNT 763 a 747 c 763 g 618 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.84e-54 Length: 2891  
Score: 611.00 Matches: 100  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 96.15% Mismatches: 2  
Query Match: 96.98% Indels: 0  
DB: 10 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AB009673 (1-2891)

QY 1 ProThrluLeuPheGluProThrluGluCysGluYasnGlyTyrValGluAlaGluGlu 20  
Db 1556 CCAACTAAGCTGTTTGAGCCACGGAATGTGAATGATATGTGAGCCGGGAGAA 1615  
QY 21 CysAspCysGlyPheHisValGluCysTyrGlyLeuGlyCysGlySerLeuSer 40  
Db 1616 TGGCACTGTGTTCCATGTGAATGCTATGAGATTGCTGAAGAGTGTGCTCTCC 1675  
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAnthSerCysLeuPheGluPro 60  
Db 1676 AATGGGCCCCATGCAATGACGCGCCCTGCTGAACAACACCTCATGCTTTTCACTCA 1735  
QY 61 ArgGlyTyrGluCysAspAlaValAlaAsnGluCysAspIleThrGluTyrCysThrGly 80  
Db 1736 CGAGGGATGATATGTCGGGATGCGGTAAACAGCTGATATCACCGAGTACTGACCTGA 1795  
QY 81 AspSerGlyGlnCysProProAsnLeuHisGlyGlnAspGlyTyrAlaCysAsnGlnAsn 100  
Db 1796 GACTCTGGCCAGTGCACCGAACCTCATTAACAAAGATGCTATGCTCAATCAAAAT 1855  
QY 101 GlnGlyArgCys 104  
Db 1856 CAGGCTGCTGC 1867

RESULT 7  
AX235014 1668 bp DNA linear PAT 11-SEP-2001  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
IMMUNEX CORPORATION (US)  
FEATURES  
source  
1. 1668  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="fusion polypeptide"  
25.1647  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC69803.1"  
/db\_xref="GI:15593674"  
/translation="MNDTLLIWLTLIIVPGSTGSGNGLVNEGECDDGFFHYECV  
LCKKSLNSGARSDDPCCNNTSCLQPRQYERDVAVEDITETGSGGCPPLV  
HKDGYACNQGRCYNGECARDNQYIWTGAAGSLFCYEKLNTGEGKNGCK  
DGRWICQSKDIVCGFLCTLNLRABRIGLOGEIIPISFYHGGRVIDSGAHVLD

DDTNGVVEDGTGPGSPNMCLDRKLIQIALNMSCLPLDSKGVCSGHVCSNEATCI  
CQFTVAGTDCISLRDPRNLAPRKDEGKPSATINRSCDTHTCPPAPAEAPSVF  
LEPPKXKDTMCSIRTEPVTVVDVSHKPEVAFNMYGVGVENAKTKPREQVNST  
YRVSVTLVHODMLNGKEYCKVSNKALPAPLEKTSIAKQOPRPOVYTLPPSDE  
LRKQVSLTCLVKGFPSPSDIAVEMESNGPENNYYKTPPVLDSDGFFLYSKULTVKS  
RMQGVNPFSCSVNHEALNHNHYOKSLSLSPGK"  
BASE COUNT 423 a 444 c 455 g 346 t  
ORIGIN

Alignment Scores:  
Pred. No.: 8.8e-52 Length: 1668  
Score: 586.00 Matches: 96  
Percent Similarity: 98.97% Conservative: 0  
Best Local Similarity: 98.97% Mismatches: 1  
Query Match: 93.02% Indels: 0  
DB: 6 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AX235014 (1-1668)

QY 8 ThrGluCysGlyYasnGlyTyrValGluAlaGluGluCysAspCysGlyPheHisVal 27  
Db 85 ACTAGTTGTGAATGATGATCGTGAAGCTGGGAGAGAGTGTATGTGTTTCAATGTG 144  
QY 28 GluCysTyrGlyLeuGlyCysGlySerLeuSerAsnGlyAlaHisCysSerAsp 47  
Db 145 GAATGCTATGATATTATGCTGAAGATGTTCCCTCCACGGGCTCACTGACAGGAC 204  
QY 48 GlyProCysCysAsnAnthSerCysLeuPheGluProAspGlyTyrGluCysAspAsp 67  
Db 205 GGGCCCTGCTGAACAATACCTCATGCTTTTTCACCCAGAGGATGAATCCGGGAT 264  
QY 68 AlaValAsnGluCysAspIleThrGluTyrCysThrGlyYasnSerGlyGlnCysProPro 87  
Db 265 GCTGTGACAGAGTGTATATTAATGATATGTAAGTGAAGACTGTGCTGCTGCCACCA 324  
QY 88 AsnLeuHisGlyGlnAspGlyTyrAlaCysAsnGlnAsnGlnGlyArgCys 104  
Db 325 AATCTTCATTAACAGACGATGATGATGATCAATCAAAATCAGCGCGCTGC 375

RESULT 8  
E10490 1464 bp RNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 1464)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Makamura, Y. and Emi, M.  
MDC PROTEIN AND DNA CODING THE SAME  
Patent: JP 1995330799-A 1 19-DEC-1995;  
JAPAN FOUND CANCER RES, EISAI CO LTD  
OS Homo sapiens (human)  
PN JP 1995330799-A/1  
PD 19-DEC-1995  
PF 22-APR-1994 JP 1994084470  
PR 14-MAY-1993 JP 93P 136602, 22-SEP-1993 JP 93P 257455, PR  
23-FEB-1994 JP 94P 49904, 12-APR-1994 JP 94P 73328 PI  
NKANURA YUSUKE, EMI MITSURU  
PC C07K16/32, C07K14/82, C12N1/21, C12N15/02, C12N15/09, C12P21/02, PC  
C12P21/08  
PC C1201/68, C01N33/53, C01N33/574, (C12N1/21, C12R1:19), (C12N15/09,  
PC C12R1:91),  
PC (C12P21/02, C12R1:19), (C12P21/08, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key Location/Qualifiers

FT source 1. 1464  
 FT /organism='Homo sapiens'  
 FT /tissue='Cerebellum'  
 FT CDS <1. .>1464  
 FT /product='MDC protein'.  
 FEATURES  
 source  
 1. 1464  
 /organism='Homo sapiens'  
 /mol\_type='genomic RNA'  
 /db\_xref='taxon:9606'  
 BASE COUNT 294 a 445 c 473 g 252 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.01e-33 Length: 1464  
 Score: 404.50 Matches: 70  
 Percent Similarity: 73.83% Conservative: 9  
 Best Local Similarity: 65.42% Mismatches: 23  
 Query Match: 64.21% Indels: 5  
 Gaps: 2  
 DB: 6  
 US-09-634-252A-4\_COPY\_496\_599 (1-104) x E10490 (1-1464)  
 QY 1 Prothrlvleuphlegluprthrglucysglvenglytrvalglualaglyglu 20  
 Db 994 CCCCTCAAGCTCTGGACCCCCCAGAGTGGGAAAGGCTTGTGGAGGCGAGGAG 1053  
 QY 21 CysAaPcysGlyPheHleValGluCysTyr-----GlyLeuCysCysLysLysCys 37  
 Db 1054 TGGCAGCTCGGCTCGGTGCAGAGTGCAGCCCGCAGAGTGCAGTCTGTCAGAAATGC 1113  
 QY 38 SerLeuSerAaNglyAlaHleCysSerAaPglyProCysCysAaAaAaThSerCysLeu 57  
 Db 1114 ACCCTGACTCAGAGCCATGTGCAGCGCGGCTCTGCTGCGCCG-----TGCAAG 1167  
 QY 58 PheGlnProAaGlyTYrgLlucysAaPgaAaValaAaNglyCysAaPglyLeuThrGluTyr 77  
 Db 1168 TACGAACCAAGGGGTGTCTCTGCGGAGAGCGCGTGAACGAGTGCAGATCGCGGAGACC 1227  
 QY 78 CysThrGlyAaPserGlyGlnCysProProAaAaLeuHleLysGlnAaPglyTYrAlaCys 97  
 Db 1228 TGCACCGGGAGCTTACGACAGTGCAGCGCGCTTAACCTGCACAACTGAGCGGTACTACTGT 1287  
 QY 98 AaNglnAaNglnGlyAaGys 104  
 Db 1288 GACCATGACAGGCGCGCTGC 1308  
 RESULT 9  
 LOCUS I25845 1464 bp DNA linear PAT 07-OCT-1996  
 DEFINITION Sequence 5 from patent US 5552526.  
 ACCESSION I25845  
 VERSION I25845.1 GI:1605715  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1464)  
 AUTHORS Nakamura, Y. and Eml, M.  
 TITLE MDC proteins and DNAs encoding the same  
 JOURNAL Patent: US 552526-A 5 03-SEP-1996;  
 FEATURES  
 source  
 1. 1464  
 /organism='unknown'  
 BASE COUNT 294 a 445 c 473 g 252 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.01e-33 Length: 1464  
 Score: 404.50 Matches: 70  
 Percent Similarity: 73.83% Conservative: 9  
 Best Local Similarity: 65.42% Mismatches: 23  
 Query Match: 64.21% Indels: 5  
 Gaps: 2

DB: 6 Gaps: 2  
 US-09-634-252A-4\_COPY\_496\_599 (1-104) x I25845 (1-1464)  
 QY 1 Prothrlvleuphlegluprthrglucysglvenglytrvalglualaglyglu 20  
 Db 994 CCCCTCAAGCTCTGGACCCCCCAGAGTGGGAAAGGCTTGTGGAGGCGAGGAG 1053  
 QY 21 CysAaPcysGlyPheHleValGluCysTyr-----GlyLeuCysCysLysLysCys 37  
 Db 1054 TGGCAGCTCGGCTCGGTGCAGAGTGCAGCCCGCAGAGTGCAGTCTGTCAGAAATGC 1113  
 QY 38 SerLeuSerAaNglyAlaHleCysSerAaPglyProCysCysAaAaAaThSerCysLeu 57  
 Db 1114 ACCCTGACTCAGAGCCATGTGCAGCGCGGCTCTGCTGCGCCG-----TGCAAG 1167  
 QY 58 PheGlnProAaGlyTYrgLlucysAaPgaAaValaAaNglyCysAaPglyLeuThrGluTyr 77  
 Db 1168 TACGAACCAAGGGGTGTCTCTGCGGAGAGCGCGTGAACGAGTGCAGATCGCGGAGACC 1227  
 QY 78 CysThrGlyAaPserGlyGlnCysProProAaAaLeuHleLysGlnAaPglyTYrAlaCys 97  
 Db 1228 TGCACCGGGAGCTTACGACAGTGCAGCGCGCTTAACCTGCACAACTGAGCGGTACTACTGT 1287  
 QY 98 AaNglnAaNglnGlyAaGys 104  
 Db 1288 GACCATGACAGGCGCGCTGC 1308  
 RESULT 10  
 LOCUS I43482 1464 bp DNA linear PAT 07-OCT-1997  
 DEFINITION Sequence 5 from patent US 5631351.  
 ACCESSION I43482  
 VERSION I43482.1 GI:2468726  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1464)  
 AUTHORS Nakamura, Y. and Eml, M.  
 TITLE Antibodies to MDC proteins  
 JOURNAL Patent: US 5631351-A 5 20-MAY-1997;  
 FEATURES  
 source  
 1. 1464  
 /organism='unknown'  
 BASE COUNT 294 a 445 c 473 g 252 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.01e-33 Length: 1464  
 Score: 404.50 Matches: 70  
 Percent Similarity: 73.83% Conservative: 9  
 Best Local Similarity: 65.42% Mismatches: 23  
 Query Match: 64.21% Indels: 5  
 Gaps: 2  
 US-09-634-252A-4\_COPY\_496\_599 (1-104) x I43482 (1-1464)  
 QY 1 Prothrlvleuphlegluprthrglucysglvenglytrvalglualaglyglu 20  
 Db 994 CCCCTCAAGCTCTGGACCCCCCAGAGTGGGAAAGGCTTGTGGAGGCGAGGAG 1053  
 QY 21 CysAaPcysGlyPheHleValGluCysTyr-----GlyLeuCysCysLysLysCys 37  
 Db 1054 TGGCAGCTCGGCTCGGTGCAGAGTGCAGCCCGCAGAGTGCAGTCTGTCAGAAATGC 1113  
 QY 38 SerLeuSerAaNglyAlaHleCysSerAaPglyProCysCysAaAaAaThSerCysLeu 57  
 Db 1114 ACCCTGACTCAGAGCCATGTGCAGCGCGGCTCTGCTGCGCCG-----TGCAAG 1167  
 QY 58 PheGlnProAaGlyTYrgLlucysAaPgaAaValaAaNglyCysAaPglyLeuThrGluTyr 77  
 Db 1168 TACGAACCAAGGGGTGTCTCTGCGGAGAGCGCGTGAACGAGTGCAGATCGCGGAGACC 1227

```

OY      78  CyeThrgIyAePSeRgIyGInCySPrOPRoAnLeuHISLySGInAspGIYTyraLaCys  97
Db      1228 TGCAccGGGgACTAGCCAGTGCcCGcCTAAGCTGcACAAGCTGGAAGGTTACTACTGT  1287
OY      98  AsnGInAnGInGIyArGcYs  104
Db      1288 GACCATGAGcAGGgCGcCTGC  1308

```

Search completed: October 21, 2003, 12:01:55  
 Job time : 3441.65 sec8





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:11:41 ; Search time 270.138 Seconds  
(without alignments)  
1039.251 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599

Perfect score: 630

Sequence: 1 PTKLEPTECGNGYVAGEE.....CPENLHKODGACNQNQGRG 104

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-DB=cn2\_1/USPTO.spool/US09634252.rnat 21102003\_090322.3362/app\_query.fasta\_1.462  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cd1  
-LIST=45 -DOCALLIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=10  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09634252.CCN 1.1 401 @runat 21102003 090322 3362 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N Geneseq 19Jun03: \*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	630	100.0	2268	21	AA245826	DNA encoding a dis
2	630	100.0	2268	22	AA511991	Human cDNA encodin
3	630	100.0	2268	24	AB078447	Nucleotide sequenc
4	630	100.0	2499	20	AA210207	CDNA encoding the
5	630	100.0	3054	20	AA278438	Human MDC3 cDNA.
6	630	100.0	3054	25	ABX76201	Lung cancer-associ
7	630	100.0	3054	25	ABX76202	Lung cancer-associ
8	630	100.0	4043	25	AA172025	ADAM 23 cDNA. Hom
9	586	93.0	1668	22	AA021438	ADAM-22dis-Fc fusi
10	567	90.0	2088	21	AA245827	Degenerate DNA enc
11	567	90.0	2088	22	AA511992	Human degenerate D
12	567	90.0	2088	24	AB078448	Degenerate nucleot
13	404.5	64.2	1464	16	AA076119	Human fetal cerebe
14	404.5	64.2	2913	16	AA076120	Human fetal cerebe
15	404.5	64.2	2923	16	AA076120	Human fetal cerebe
16	404.5	64.2	3183	16	AA076122	Human fetal brain
17	363.5	57.7	2604	20	AA210208	CDNA encoding beta
18	363.5	57.7	2697	20	AA278436	Human MDC2-beta cD
19	363.5	57.7	2805	20	AA278437	Human MDC2-alpha c
20	338.5	53.7	1674	22	AA021438	ADAM-22dis-Fc fusi
21	292	46.3	2642	24	AB086500	Human novel protei
22	292	46.3	2649	24	AA597181	Human novel protei
23	292	46.3	2705	24	AB084499	Human metallopept
24	292	46.3	2762	24	AB084498	Human novel protei
25	292	46.3	3431	24	AA085513	First splice varia
26	292	46.3	3431	24	AB072136	Nucleotide sequenc
27	292	46.3	3446	24	AB066136	CDNA encoding huma
28	292	46.3	3468	22	AA085514	Second splice vari
29	292	46.3	3468	24	AB072137	Nucleotide sequenc
30	292	46.3	3509	23	AB072025	Gene 216 alternati
31	292	46.3	3509	25	ABX74890	Human gene 216 cDN
32	292	46.3	3562	22	AA046157	Human DNA encoding
33	292	46.3	3562	25	AA046157	Human PRO1891 cDNA
34	292	46.3	3562	25	ABX98185	Human cDNA encodin
35	292	46.3	3562	25	ABX98185	Human cDNA encodin
36	292	46.3	3562	25	ABX98187	Human secreted/tra
37	292	46.3	3562	25	ABX97976	Human PRO polynuc
38	292	46.3	3562	25	ABX78760	Human PRO polynuc
39	292	46.3	3562	25	ABX75773	Human cDNA encodin
40	292	46.3	3562	25	ABX76878	Human PRO polynuc
41	292	46.3	3562	25	ABX16818	Human cDNA encodin
42	292	46.3	3626	23	AB072023	Gene 216 encoding
43	292	46.3	3626	25	ABX74888	CDNA encoding huma
44	279	44.3	1638	22	AA021434	ADAM-15dis-Fc fusi
45	278	44.1	2653	17	AA034616	CRII-7 nerve prote

## ALIGNMENTS

RESULT 1  
AA245826  
ID AA245826 standard; DNA: 2268 BP.  
XX  
AC AA245826;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE DNA encoding a disintegrin homologue designated zdint1.  
XX  
KW Human: disintegrin homologue; zdint1; cardiac myocyte; adipocyte;  
KW gene therapy; cell-cell interaction; chondro sarcoma; atherosclerosis;  
KW Alzheimer's disease; resenosis; ischemic reperfusion; obesity;  
KW intimal hyperplasia; tumour; platelet aggregation; apoptosis;  
KW neurogenesis; myogenesis; arthritis; myogenesis; neurogenesis;  
KW connective tissue disorder; chondrogenesis; tumour proliferation;  
KW inflammation; ss.

```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 3..2093
XX FT /tag= a
XX FT /product= "zdint1"
XX PN WO200002912-A2.
XX PD 20-JAN-2000.
XX PE 09-JUL-1999; 99WO-US15638.
XX PR 10-JUL-1998; 98US-0113883.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Sheppard PO, Baidur N, Deisher TA, Bishop PD;
XX DR WPI; 2000-160898/14.
XX DR P-PSDB; AAU54457.
XX PT Polypeptide useful in modulating cell-cell interaction in tissues of
XX PT heart, brain, spinal cord and treating chondro sarcoma,
XX PT atherosclerosis, restenosis, obesity, intimal hyperplasia and tumors -
XX PS Example 1; Page 123-127; 132pp; English.
XX
CC The present sequence encodes a human disintegrin homologue, designated
CC zdint1. The zdint1 polypeptide is a cardiac myocyte proliferation and
CC differentiation stimulator, as well as an adipocyte proliferation and
CC differentiation inhibitor. Polynucleotides encoding zdint1 are used in
CC gene therapy. The zdint1 polypeptide is useful in modulating cell-cell
CC interactions of cells derived from tissues of heart, brain, spinal cord
CC and skeletal muscle. It is useful in treating and diagnosing chondro
CC sacroma, atherosclerosis, Alzheimer's disease, restenosis, ischemic
CC reperfusion, obesity, intimal hyperplasia and tumors of heart, brain,
CC and spinal cord. The zdint1 polypeptide is also useful in identifying
CC its new family members, antagonists, agonists and antibodies.
CC Antagonists, antibodies and fusion proteins of zdint1 are useful in
CC inhibiting platelet aggregation, apoptosis, neurogenesis and myogenesis.
CC Agonists and antagonists are useful in studying cell-cell interactions,
CC arthritis, myogenesis, neurogenesis, connective tissue disorders,
CC chondrogenesis, tumour proliferation and suppression, extracellular
CC matrix proteins, repair and remodelling of ischemic reperfusion,
CC inflammation, and apoptosis.
XX SO Sequence 2268 BP; 654 A; 457 C; 558 G; 598 T; 1 other;
XX
XX Alignment Scores:
XX Pred. No.: 3..496-55 Length: 2268
XX Score: 630.00 Matches: 104
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-634-252a-4_COPY_496_599 (1-104) x AA245826 (1-2268)
XX
XX QY 1 PROTHIRYSLNLEPHGLUPROTHIRYSLNGLYTYRVAIGLNAAGLYGLU 20
XX DB 1119 CCMAAAGGCTATTGAGCCCGGCAATGTGTAATGCTGAGAGCTGGAGAG 1178
XX
XX QY 21 CYSASPQYSGLYPHEHISVAIGLUCYSTRYGLYLUCYSCYSLVLYSCYSSERLEUS 40
XX DB 1179 TGTGATTTGTGTTTCAATGTAATGCTATGATGCTGTAAGAAATGTTCCCTCTCC 1238
XX
XX QY 41 ASNGIYLAHISCYSSERASPGIYPROCYSYSAASANTHSECYSLNPHGLNPRO 60
XX DB 1239 AACGGGGCTCACTGACGACGACGGCCCTGCTTAACATACCTCATGCTTTTACGCA 1298
XX
XX QY 61 ARGGLYTYRGUCYARGAPALAVALSNGIUCYASPIETHTHGLUTYRCYATHRG 80
```

```
DB 1299 CGAGGGTATGATCCCGGAGTCTGTGAACAGCTGATATTAAGTAATGTA 1358
QY 81 ASPERGILYGLNICYSPROPROASNLEWHISYGLINSPGIYTYRAACYSASGLN 100
DB 1359 GACTCTGTCAGTCCACCAAAATCTTCAATAGCAAGACGATATGCAATCAAAAT 1418
QY 101 GINGIYARGCY 104
DB 1419 CAGGGCCGCTGC 1430
XX
XX RESULT 2
XX AAS11991
XX ID AAS11991 standard; CDNA; 2268 BP.
XX AC AAS11991;
XX DT 04-DEC-2001 (first entry)
XX DE Human cDNA encoding partial disintegrin protease zdint1.
XX
XX Human; disintegrin; zdint1; ss; antiangiogenic; vascular; thrombolytic.
XX cell matrix; cell-cell interactions; apoptosis; neurogenesis;
XX connective tissue disorders; chondrogenesis; arthritis;
XX tumour proliferation; ischaemia reperfusion; inflammation;
XX chromosome 2q33.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 3..2093
XX FT /tag= a
XX FT /product= "zdint1"
XX FT /partial
XX FT /note= "No start codon"
XX FT sig_peptide 3..491
XX FT /tag= b
XX FT /note= "Encodes propeptide sequence"
XX FT mat_peptide 492..2090
XX FT /tag= c
XX FT /label= Mature_zdint1
XX
XX US6265199-B1.
XX 24-JUL-2001.
XX PD 09-JUL-1999; 99US-0351414.
XX PR 10-JUL-1998; 98US-0092371.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Sheppard PO, Baidur N, Deisher TA, Bishop PD, Taft DW;
XX DR WPI; 2001-450736/48.
XX DR P-PSDB; AAU07190.
XX
XX Disintegrin protease zdint1, useful for producing agents for the
XX development of antithrombotic and anti-migration of tumour cells and
XX have antiangiogenic activity -
XX
XX Example 1; Column 47-53; 50pp; English.
XX
XX The invention relates to an isolated novel disintegrin protease family
XX member, zdint1. Disintegrins bind cell surface molecules, including
XX integrins, on the surface of various cells such as platelets,
XX fibroblasts, tumour, endothelial, muscle, neuronal, bone and sperm cells.
XX Disintegrins are unique and potentially useful tools for investigating
XX cell matrix and cell-cell interactions, apoptosis, neurogenesis,
XX connective tissue disorders, chondrogenesis, arthritis, tumour
XX proliferation, ischaemia reperfusion and inflammation. Additionally, they
XX are useful in the development of antithrombotic and anti-migration of
XX tumour cells and have antiangiogenic activity. The present sequence
```

CC encodes human disintegrin, zdn1, the gene for which is located on  
XX chromosome 2q33.  
SQ Sequence 2268 BP; 654 A; 457 C; 558 G; 598 T; 1 other;

## Alignment Scores:

Pred. No.:	3,496-55	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AAS11991 (1-2268)

QY 1 ProThrlsLeuphegIupProThrgIuCySgIyASnGlyTYrYValGluAglYgluGlu 20  
DB 1119 CCAACAAAGCTATTGAGCCACGGAATGTGAAATGATACGTGAAAGCTGGGAGAG 1178  
QY 21 CysAspCysGlyPheHisValGluCyTYrGlyLeuCySgIySlyblyScySerLeuSer 40  
DB 1179 TGTGATTGTGGTTTTCATGTGAAATGCTATGATGATTATGCTGTAAATGTTCCCTCTCC 1238  
QY 41 AsnGlyAlaHisCysSerAspGlyProCySgAsnAsnThrSerCysLeuphegIupPro 60  
DB 1239 AACGGGCTCAGTCAGCAGCGAGCGGCTGCTGTAAACATACCTCTTTTTCACCA 1298  
QY 61 ArgGlyTYrGluCySgAspAlaValAsnGluCySgIySlyblyScySerLeuSer 80  
DB 1299 CGAGGATGAAATGCGCGGATGCTGTAAACAGTGTATTTACTGAAATGTGACTGGA 1358  
QY 81 AspSerGlyGlnCySPProAsnLeuHisGlySgIySlyblyScySerLeuSer 100  
DB 1359 GACTCTGTAGTCCACCAACAACTTCTATAGCAACAGGATNGCATGCAATCAAAAT 1418  
QY 101 GlnGlyArgCys 104  
DB 1419 CAGGCGCGCTGC 1430

## RESULT 3

ABO78447 ID ABO78447 standard; cDNA; 2268 BP.

XX AC ABO78447;

DT 05-NOV-2002 (first entry)

XX Nucleotide sequence of human zdn1.

XX Human; zdn1; disintegrin protease; platelet accumulation;  
XX chromosome 2q33; platelet aggregation; proteolysis; apoptosis;  
XX neurogenesis; myogenesis; connective tissue disorder; arthritis;  
XX chondrogenesis; cell adhesion; cell fusion; neuron; myocyte; heart;  
XX brain; spinal cord; skeletal muscle tissue; Alzheimer's disease;  
XX tumour formation; multiple sclerosis; congestive heart failure;  
XX ischaemic reperfusion; intimal hyperplasia; restenosis; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 3..2093

XX FT /\*tag= a

XX FT /product= "zdn1"

XX PN US2002072102-A1.

XX PD 13-JUN-2002.

XX PF 16-MAR-2001; 2001US-0809790.

XX PR 10-JUL-1998; 98US-092371P.

XX PR 09-JUL-1999; 99US-0351414.

XX XX

PA (SHEP/) SHEPPARD P O.  
PA (BAIN/) BAINDUR N.  
PA (DEIS/) DEISHER T A.  
PA (BISH/) BISHOP P D.

PI Sheppard PO, Baidur N, Deisher TA, Bishop PD;

XX WPI; 2002-598452/64.

DR P-PSDB; ABB78130.

PT New disintegrin homolog polypeptide and polynucleotide, useful for

PT modulating cell-cell interactions and diagnosis, treatment of  
PT Alzheimer's disease, tumour formation, multiple sclerosis and congestive  
PT heart failure

PS Claim 14; Page 26-29; 53pp; English.

CC The present sequence encodes a human polypeptide designated zdn1.  
CC zdn1 is a member of the disintegrin protease family. zdn1 inhibits  
CC platelet accumulation. The zdn1 gene is present on chromosome 2q33.  
CC zdn1 polypeptides and polynucleotides are useful in treatment of  
CC disorders associated with infarct in brain or heart tissue and/or  
CC platelet aggregation, to modulate proteolysis, apoptosis, neurogenesis,  
CC myogenesis, connective tissue disorders, arthritis, chondrogenesis,  
CC cell adhesion, cell fusion, and signalling or to treat or prevent  
CC development of pathological conditions in such diverse tissue as heart,  
CC brain, spinal cord and skeletal muscle. The molecules modulate  
CC inhibition and proliferation of neurons and myocytes in heart, brain,  
CC spinal cord and skeletal muscle tissue. Disorders which may be amenable  
CC to diagnosis, treatment or prevention with zdn1 polypeptides include,  
CC Alzheimer's disease, tumour formation, multiple sclerosis, congestive  
CC heart failure, ischaemic reperfusion or infarct and degenerative  
CC diseases. The zdn1 molecules particularly useful in the treatment of  
CC intimal hyperplasia or restenosis due to acute vascular injury.

XX SQ Sequence 2268 BP; 654 A; 457 C; 558 G; 598 T; 1 other;

## Alignment Scores:

Pred. No.:	3,496-55	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x ABO78447 (1-2268)

QY 1 ProThrlsLeuphegIupProThrgIuCySgIyASnGlyTYrYValGluAglYgluGlu 20  
DB 1119 CCAACAAAGCTATTGAGCCACGGAATGTGAAATGATACGTGAAAGCTGGGAGAG 1178  
QY 21 CysAspCysGlyPheHisValGluCyTYrGlyLeuCySgIySlyblyScySerLeuSer 40  
DB 1179 TGTGATTGTGGTTTTCATGTGAAATGCTATGATGATTATGCTGTAAATGTTCCCTCTCC 1238  
QY 41 AsnGlyAlaHisCysSerAspGlyProCySgAsnAsnThrSerCysLeuphegIupPro 60  
DB 1239 AACGGGCTCAGTCAGCAGCGAGCGGCTGCTGTAAACATACCTCTTTTTCACCA 1298  
QY 61 ArgGlyTYrGluCySgAspAlaValAsnGluCySgIySlyblyScySerLeuSer 80  
DB 1299 CGAGGATGAAATGCGCGGATGCTGTAAACAGTGTATTTACTGAAATGTGACTGGA 1358  
QY 81 AspSerGlyGlnCySPProAsnLeuHisGlySgIySlyblyScySerLeuSer 100  
DB 1359 GACTCTGTAGTCCACCAACAACTTCTATAGCAACAGGATNGCATGCAATCAAAAT 1418  
QY 101 GlnGlyArgCys 104  
DB 1419 CAGGCGCGCTGC 1430

## RESULT 4

AA210207

ID	AA10207 standard; DNA; 2499 BP.
XX	AA10207 standard; DNA; 2499 BP.
AC	AA10207;
XX	
DT	29-OCT-1999 (first entry)
XX	
DE	cDNA encoding the human SVPH3-17 protein.
XX	
KW	SVPH3-17; metalloproteinase-disintegrin; human chromosome 2;
KW	human chromosome 7; malignant hyperthermia susceptibility;
KW	Zellweger syndrome; neonatal adrenoleukodystrophy;
KW	infantile Refsum disease; progressive familial intrahepatic cholestasis;
KW	mucopolysaccharidosis VII; split hand/foot malformation;
KW	archoylomogenic right ventricular displasia-4; Coppock-like cataract;
KW	insulin dependent diabetes mellitus-12; lamellar type ichthyosis;
KW	transient neonatal myasthenia gravis; congenital aculeiform cataract;
KW	juvenile amyotrophic lateral sclerosis;
KW	familial paroxysmal choreoathetosis;
KW	Finnish lethal neonatal metabolic syndrome; ss.
XX	
OS	Homo sapiens.
XX	
XX	Key
FT	Location/Qualifiers
FT	CDS
FT	1..2499
FT	/*tag= a
XX	
PN	W09941388-A2.
XX	
PD	19-AUG-1999.
XX	
PF	11-FEB-1999; 99WO-US03016.
PR	11-FEB-1998; 98US-0074310.
XX	
PA	(IMMV ) IMMUNEX CORP.
PI	Cerretci DP;
PI	
XX	
DR	WPI; 1999-527371/44.
DR	P-PSDB; AA930207.
XX	
PT	DNA encoding the SVPH3-13 and SVPH3-17 proteins for detecting
PT	disease corresponding to chromosome 7, e.g. Zellweger syndrome
XX	
XX	Claim 1; Page 7-8; 82pp; English.
XX	
CC	The present sequence encodes a protein designated SVPH3-17, which
CC	is a member of the metalloproteinase-disintegrin family. The
CC	specification also describes SVPH3-13 proteins. Both sequences can
CC	be used to identify human chromosome 2 or 7, and to map genes on these
CC	two chromosomes, and also to identify genes associated with certain
CC	diseases, syndromes, or other human conditions associated with human
CC	chromosome 2 or 7. The disease that correspond to chromosome 7 include
CC	malignant hyperthermia susceptibility, Zellweger syndrome, neonatal
CC	adrenoleukodystrophy, infantile Refsum disease, progressive familial
CC	intrahepatic cholestasis, mucopolysaccharidosis VII, and split
CC	hand/foot malformation. Diseases associated with chromosome 2 include
CC	archoylomogenic right ventricular displasia-4, insulin dependent diabetes
CC	mellitus-12, transient neonatal myasthenia gravis, juvenile amyotrophic
CC	lateral sclerosis, congenital aculeiform cataract, Coppock-like cataract,
CC	lamellar type ichthyosis, familial paroxysmal choreoathetosis, and
CC	Finnish lethal neonatal metabolic syndrome.
XX	
XX	
SO	Sequence 2499 BP; 677 A; 593 C; 645 G; 584 T; 0 other;

Alignment Scores:

Pred. No.:	3,95e-55	Length:	2499
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

```

US-09-634-252A--4_COPY_496_599 (1-1-104) x AAZ10207 (1-24-99)

Oy 1 ProThrlYslEupheGluPProThrlGluCySGlYAsnGlyTYrYalGluAlaGlyGluGlu 20
Db 1486 CCAACAAAGCTATTGTGAGCCACGCAAGTGTGAATGATACGTGAAGCTGGGGAGAG 1545
Oy 21 CysAspCyAGlYPhEhISa1GluCySTyrGlyLeuCYcSylSLySCysSerLeuSer 40
Db 1546 TGTGATTGTGTGTTTTCATGTGGAAATGCTATGGAATTATGCTGTAAAGAAATGTTCCCTTCC 1605
Oy 41 AsnGlyAlaHisCysSerAspGlyProCYcCysAsnAsnThrSerCYcLeuPheGlnPro 60
Db 1606 AAGCGGGCTCACTGCGACGACGACGGCCCTGCTGTAACAATACCTCATCTCTTTTCAGCCA 1665
Oy 61 ArgGlyTYrGluCYcAspAGsPAlaValAsnGluCYcAsp11ethrGluTYrCYcThGly 80
Db 1666 CGAGGGTATGAATGCCGGGATGCTGTGAACAGAGTGATATTACTGAATATTGTACTGGA 1725
Oy 81 AspSerGlyGlnCYcProProlAsnLeuHisGlyGlnAspGlyTYrAlaCYcAsnGlnAsn 100
Db 1726 GACTCTGTGTCAGTCCACCAAAATCTTCATPAAGACAGATATGCAATGCAATCAAAAT 1785
Oy 101 GlnGlyArgCYc 104
Db 1786 CAGGGCCGCTGC 1797

RESULT 5
AAx78438 standard; cDNA to mRNA; 3054 BP.
ID AAx78438
XX AAx78438;
AC AAx78438;
XX 26-AUG-1999 (first entry)
DT
XX
DE Human MDC3 cDNA.
XX
KM Metalloproteinase-like-disintegrin-like cysteine rich protein; human;
KW MDC2-alpha; MDC2-beta; MDC3; medical; treatment; diagnosis; ds.
XX
OS Homo sapiens.
XX
XX
XX Key location/Qualifiers
XX FH 224..2722
XX FT /*tag= a
XX FT /product= "MDC3"
XX
XX JF1155574-A.
XX
XX 15-JUN-1999.
XX
XX 01-DEC-1997; 97JP-0330020.
XX
XX 01-DEC-1997; 97JP-0330020.
XX
XX (EISA ) EISA1 CO LTD.
XX
XX WPI; 1999-398071/34.
XX DR P-PSDB; AAY25120.
XX
XX
XX New protein belonging to MDC gene family - useful in medical
XX treatment and diagnosis
XX
XX Claim 17; Page 13-16; 17pp; Japanese.
XX
XX
XX This invention describes novel human MDC2-alpha, MDC3-beta and MDC3
XX proteins and their encoding nucleic acids. The products of the invention
XX are useful in medical treatment and diagnosis.
XX
XX Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other;

Alignment Scores:
Pred. No.: 5,08e-55 Length: 3054
Score: 630.00 Matches: 104

```

Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 20  
Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AAX78438 (1-3054)

QY 1 ProThrlYsLeuPheGluProThrGluCyseGlyAasnGlyTyryValGluAglYglUglU 20  
Db 1709 CCAACAAAGCTATTGGAGCCACCGGAATGCGAATGATACGTGGAAGCTGGGAGAG 1768  
QY 21 CysAspCyseGlyPheHisValGluCyseTyryGlyLeuCyseGlySlySlyCyseSerLeuSer 40  
Db 1769 TGTATTGTGGTTTTCATGATGGAATGCTATGCTGTAAAGAAATGTTCCCTCTCC 1828  
QY 41 AasnGlyAHisCyseSerAspGlyProCyseCysAasnAnThrSerCyseLeuPheGluPro 60  
Db 1829 AACGGGGCTCACTGACGAGCGAGGGCCCTGCTGTAAACATACCTCTGCTTTTCAGCCA 1888  
QY 61 ArgGlyTyryGluCyseArgAspAlaValAasnGluCyseAspIleThrGluTyryCyseThrGly 80  
Db 1889 CGAGGATGATGAATGCCGGATGCTGTGAACGAGTGTATTTACTGATATTGTACTGGA 1948  
QY 81 AspSerGlyGlnCyseProProAsnLeuHisGlyGlnAspGlyTyryAlaCyseAasnGlnAsn 100  
Db 1949 GACTCTGGTCAGTGCACCAACAAATCTTCATAAAGCAAGCATATGCAATCAAAAT 2008  
QY 101 GlnGlyAArgCys 104  
Db 2009 CAGGGCCGCTGC 2020

## RESULT 6

ABX76201  
ID ABX76201 standard; DNA; 3054 BP.

AC ABX76201;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #72.

KW Lung cancer-associated polynucleotide; gene; ds; cytotatic; emphysema;  
antiinflammatory; antisthmatic; non-small cell lung cancer; atelectasis;  
small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

PN MO200286443-A2.

PD 31-OCT-2002.

PF 18-APR-2002; 2002WO-US12476.

PR 18-APR-2001; 2001US-284770P.

PR 10-MAY-2001; 2001US-290492P.

PR 09-NOV-2001; 2001US-339245P.

PR 13-NOV-2001; 2001US-350666P.

PR 29-NOV-2001; 2001US-334370P.

PR 12-APR-2002; 2002US-372246P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Aziz N, Murray R;

DR WPI, 2003-093161/08.

DR P-PSDB; ABUS6479.

XX

XX Claim 22; Page 243; 453pp; English.  
PS The invention relates to a method for detecting a lung cancer-associated  
XX transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung  
CC cancer-associated polynucleotides and polypeptides are used for  
CC identifying a compound that modulates a lung cancer-associated  
CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
CC cell to treat lung cancer in a patient and for treating a mammal having  
CC lung cancer by administering a modulatory compound identified. The  
CC methods are useful for treating lung cancer, such as small cell lung  
CC cancer, non-small cell lung cancer or other benign or precancerous  
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
CC and polypeptides are useful for diagnostic purposes and as targets for  
CC screening for therapeutic compounds that modulate lung cancer, such as  
CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated  
CC polynucleotides of the invention.

XX Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other:  
Alignment Scores:  
Pred. No.: 5,08e-55 Length: 3054  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x ABX76201 (1-3054)

QY 1 ProThrlYsLeuPheGluProThrGluCyseGlyAasnGlyTyryValGluAglYglUglU 20  
Db 1709 CCAACAAAGCTATTGGAGCCACCGGAATGCGAATGATACGTGGAAGCTGGGAGAG 1768  
QY 21 CysAspCyseGlyPheHisValGluCyseTyryGlyLeuCyseGlySlySlyCyseSerLeuSer 40  
Db 1769 TGTATTGTGGTTTTCATGATGGAATGCTATGCTGTAAAGAAATGTTCCCTCTCC 1828  
QY 41 AasnGlyAHisCyseSerAspGlyProCyseCysAasnAnThrSerCyseLeuPheGluPro 60  
Db 1829 AACGGGGCTCACTGACGAGCGAGGGCCCTGCTGTAAACATACCTCTGCTTTTCAGCCA 1888  
QY 61 ArgGlyTyryGluCyseArgAspAlaValAasnGluCyseAspIleThrGluTyryCyseThrGly 80  
Db 1889 CGAGGATGATGAATGCCGGATGCTGTGAACGAGTGTATTTACTGATATTGTACTGGA 1948  
QY 81 AspSerGlyGlnCyseProProAsnLeuHisGlyGlnAspGlyTyryAlaCyseAasnGlnAsn 100  
Db 1949 GACTCTGGTCAGTGCACCAACAAATCTTCATAAAGCAAGCATATGCAATCAAAAT 2008  
QY 101 GlnGlyAArgCys 104  
Db 2009 CAGGGCCGCTGC 2020

## RESULT 7

ABX76292  
ID ABX76292 standard; DNA; 3054 BP.

AC ABX76292;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #156.

KW Lung cancer-associated polynucleotide; gene; ds; cytotatic; emphysema;  
antiinflammatory; antisthmatic; non-small cell lung cancer; atelectasis;  
small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

MO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US12476.

18-APR-2001; 2001US-284770P.

10-MAY-2001; 2001US-290492P.

09-NOV-2001; 2001US-339245P.

13-NOV-2001; 2001US-350666P.

29-NOV-2001; 2001US-334370P.

12-APR-2002; 2002US-372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Aziz N, Murray R;

WPI; 2003-093161/08.

P-PSDB; ABUS6563.

Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.

Claim 22; Page 307-308; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention.

Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other;

Alignment Scores:

Pred. No.:	5,08e-55	Length:	3054
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x ABX76292 (1-3054)

ProthrlysluPhuGluProthrGluCysGlyAenGlyTYrValGluAlaGluGlu 20

1709 CCAACAAAGCTATTTGAGCCCGGAAATGTGAATGATGAGTGGAGGAGG 1768

21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysCysLeuLysCysSerLeuSer 40

1769 TGTGATTTGGTTTCATGTGGAATGCTATGATTAATGTAAGAAATGTTCCCTCTCC 1828

41 AenGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGluPro 60

1829 AACGGGCTCACTGACGAGCGGCCCTGCTGTAAACAATACCTGCTTTTCAGCCA 1888

61 ArgGlyTYrGluCysArgAspAlaValAenGluCysAspIleThrGluTYrCysThrGly 80

1889 CGAGGATATGAATGCCGGATGCTGTGAACAGTGTGATATTACTGAATATTGTACTGA 1948

81 AspSerGlyGlnCysProProAsnLeuHisIysGlnAspGlyTYrAlaCysAsnGlnAsn 100

1949 GACTCTGTCAGTGTGCCCAACCAATCTTCATTAAGCAAGCATATGATCAATCAAAAT 2008

101 GlnGlyArgCys 104

2009 CAGGCCCGCTGC 2020

RESULT 8

AA172025

AA172025 standard; cDNA, 4043 BP.

AA172025;

04-MAR-2002 (first entry)

ADAM 23 cDNA.

Disintegrin; ADAM 23; metalloproteinase; domain; adhesion; proteolysis; modulator; alpha-v-beta3 integrin; tumour progression; neural tissue; angiogenesis; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 1017..3515

FT /\*tag= a

FT /product= "ADAM 23"

MO200174857-A2.

11-OCT-2001.

02-APR-2001; 2001WO-US10729.

03-APR-2000; 2000US-194164P.

(BRIM ) BRISTOL-MYERS SQUIBB CO.

Lopez-Otin C, Freije JMP, Bianchi AB, Miguel SC, Garcia JML;

Trail P;

WPI; 2002-066298/09.

P-PSDB; AAB47778.

Nucleic acid encoding a new ADAM family member, designated ADAM 23 is useful to find modulators of its interaction with integrin which can be used to prevent angiogenesis or increase neural growth.

Claim 2; Page 36-38; 44pp; English.

This sequence encodes a cellular disintegrin, ADAM 23, which is a protein having a disintegrin and metalloproteinase domain. This protein can perform both adhesion and proteolytic functions. Modulators of the interaction between ADAM and alpha-v-beta3 integrin are used to inhibit tumour progression or induce growth of neural tissue. These compounds modulate angiogenesis and induction of matrix metalloproteinases facilitating migration of tumour cells and growth of neural tissue.

Sequence 4043 BP; 1110 A; 946 C; 1002 G; 985 T; 0 other;

Alignment Scores:

Pred. No.:	7.21e-55	Length:	4043
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AA172025 (1-4043)

QY 1 Prothrtysleupheglnuprothrglucysgllysennglytyrvalgluaaglglglu 20  
 DB 2502 CCACCAAGACTATTGAGCCACCGAATGTGGAATGATACGTAAGAAATGTTCCCTCTCC 2561  
 QY 21 CysAspCysGlyPheHisValGlcysGlyTgLyLeuGlyCysGlySLysCysSerLeuSer 40  
 DB 2562 TGTCATTGTGTTTTCATGTGCAATGCTATGAGATTATGCTTAAGAAATGTTCCCTCTCC 2621  
 QY 41 AsnGlyAlahisCysSerAspGlyProCysCysAsnAnthrSerCysLeuPheGlnPro 60  
 DB 2622 AACGGGCTCACTCAGAGAGAGGGCCCTGCTGTAACATACCTCATGCTTTTCACCA 2681  
 QY 61 ArgGlyTyrGlyCysArgAspAlaValAsnGlyCysAspIlethrglyTyrCysThgLy 80  
 DB 2682 CGAGGGTATGATGCCCGGATGCTGTAAACAGAGTGTATATCTGATATGACTGGA 2741  
 QY 81 AspserGlyGlnCysProProAsnLeuHislyGlnAspGlyTyrAlaCysAsnGlnAsn 100  
 DB 2742 GACTCTGTGTCAGTCCCAAACTTTCATAGCAAGCGATATGCAATCAAAAT 2801  
 QY 101 GlnGlyArgCys 104  
 DB 2802 CAGGGCCGCTGC 2813

RESULT 9  
 AAD21439 standard; DNA; 1668 BP.

AC AAD21439:  
 XX 28-JAN-2002 (first entry)  
 XX  
 DE ADAM-23dis-Fc fusion construct DNA.  
 XX  
 KM Human; ADAM disintegrin domain; integrin; endothelial cell migration;  
 KM angiogenesis; ocular disorder; inflammatory disease; bone resorption;  
 KM osteoporosis; restenosis; thrombosis; tissue repair; wound healing;  
 KM retinopathy; retinodisplasia; neovascular glaucoma; macular degeneration;  
 KM retrolental fibroplasia; inflammatory bowel disease; rubecosis; uveitis;  
 KM arthritis; rheumatism; myocardial infarction; coronary artery disease;  
 KM tumour; stroke; atherosclerosis; arteriosclerosis; injury; psoriasis;  
 KM preclampsia; embolism; ischaemia; fusion construct; immunoglobulin K;  
 KM Igk; Fc region; antiinflammatory; osteopathic; vasotropic; thrombolytic;  
 KM  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 25..1647  
 FT /\*tag= a  
 FT /\*product= "ADAM-23dis-Fc fusion construct"  
 FT sig\_peptide 25..84  
 FT /\*tag= b  
 FT /\*note= "Igk leader sequence"  
 FT mat\_peptide 85..1644  
 FT /\*tag= c  
 FT /\*product= "Mature ADAM-23dis-Fc fusion construct"  
 FT 93..954  
 FT /\*tag= d  
 FT /\*note= "Human ADAM disintegrin"  
 FT misc\_feature 963..1644  
 FT /\*tag= e  
 FT /\*note= "Fc region"  
 XX  
 PN WO200162905-A2.  
 XX  
 PD 30-AUG-2001.

XX 23-FEB-2001; 2001WO-US05701.  
 PF  
 XX  
 PR 25-FEB-2000; 2000US-184865P.  
 XX  
 PA (IMMUNEX CORP.  
 PI Fanslow WC, Cerretti DP, Poindexter KM, Black RA;  
 DR WPI: 2001-625725/72.  
 DR P-PDSB: AAE13059.  
 XX  
 PT Antagonizing the binding of an integrin to its ligand useful for the  
 PT treatment of angiogenesis comprises administration of an  
 PT ADAM-disintegrin domain polypeptide  
 XX  
 PS Claim 14; Page 58-60; 66pp; English.  
 XX  
 CC The invention relates to the method and use of ADAM disintegrin domain  
 CC polypeptides for inhibiting the biological activity of integrins,  
 CC endothelial cell migration and angiogenesis. ADAM disintegrin domain  
 CC polypeptides are used for treatment of ocular disorders, malignant and  
 CC metastatic conditions, inflammatory diseases, osteoporosis and other  
 CC conditions mediated by accelerated bone resorption, restenosis,  
 CC inappropriate platelet activation, recruitment or aggregation, thrombosis  
 CC or a condition requiring aggregation, thrombosis or a condition requiring  
 CC tissue repair or wound healing, angiogenesis, ocular neovascularisation  
 CC or solid tumour, diabetic retinopathy, retinopathy or prematurity,  
 CC neovascular glaucoma, retinodisplasia, retrolental fibroplasia, rubecosis,  
 CC uveitis, macular degeneration and corneal graft neovascularisation.  
 CC inflammatory diseases, ocular tumours, diseases associated with choroidal  
 CC or iris neovascularisation, arthritis, rheumatism, inflammatory bowel  
 CC disease, psoriasis, coronary artery disease or injury, myocardial  
 CC infarction or injury following myocardial infarction, stroke, unstable  
 CC angina, atherosclerosis, arteriosclerosis, preclampsia, embolism,  
 CC platelet-associated ischaemic disorders including lung ischaemia,  
 CC coronary ischaemia, cerebral ischaemia, restenosis following percutaneous  
 CC coronary intervention including angioplasty, atherectomy, stent placement  
 CC and bypass surgery, thrombotic disorders including coronary artery  
 CC thrombosis, cerebral artery thrombosis, intracardiac thrombosis,  
 CC peripheral artery thrombosis, venous thrombosis, thrombosis and  
 CC coagulopathies associated with exposure to a foreign or injured tissue  
 CC surface and reocclusion following thrombosis, deep venous thrombosis,  
 CC pulmonary embolism, transient ischaemic attacks and another conditions  
 CC where vascular occlusion is a common underlying feature. In individuals  
 CC at high risk for thrombus formation of reformation, advanced coronary  
 CC artery disease, or for occlusion, reocclusion, stenosis and/or restenosis  
 CC of blood vessels or stroke benign tumours and preneoplastic conditions,  
 CC myocardial angiogenesis, haemophilic joints, scleroderma, vascular  
 CC adhesions, asthma and allergy, eczema and dermatitis, graft versus host  
 CC disease, sepsis, adult respiratory distress syndrome, telangiectasia and  
 CC wound granulation. The method are used in combination with angioplasty  
 CC procedures, such as balloon angioplasty, laser angioplasty, coronary  
 CC atherectomy or similar techniques, carotid endarterectomy, anastomosis of  
 CC vascular grafts, surgery having a high risk of thrombus formation (i.e.  
 CC coronary bypass surgery, insertion of a prosthetic valve or vessel and  
 CC the like), atherectomy, stent placement, placement of a chronic  
 CC cardiovascular device such as an in-dwelling catheter or prosthetic valve  
 CC or vessel, organ transplantation or bypass surgery. The present sequence  
 CC is a DNA encoding ADAM disintegrin domain polypeptide fusion construct.  
 CC The fusion construct comprises of immunoglobulin K (Igk) leader, human  
 CC ADAM disintegrin and Fc region.  
 XX  
 SQ Sequence 1668 BP; 423 A; 444 C; 455 G; 346 T; 0 other;

Alignment Scores:  
 Pred. No.: 8, 15e-51 Length: 1668  
 Score: 586.00 Matches: 96  
 Percent Similarity: 98.97% Conservative: 0  
 Best Local Similarity: 98.97% Mismatches: 1  
 Query Match: 93.02% Indels: 0  
 DB: 22 Gaps: 0

US-09-634-252a-4\_copy\_496\_599 (1-104) x AAD21439 (1-1668)

QY 8 ThrGluCysGlyAsnGlyTyrValGluAlaGluGluCysAspCysGlyPheHisVal 27  
DB 85 ACTACTGTGGAAAGCATACCTCCAGAGCTGGGAGAGACTGATGTTGCTTTTATGTG 144

QY 28 GluCyTYrGlyLeuCyCysGlySylsCysSerLeuSerAsnGlyAlaHisCysSerAsp 47  
DB 145 GAATGCTATGATATGCTGTGAAGAAATGTTCCCTCTCCAAAGGGGCTCAGCAGCGAC 204

QY 48 GlyProCysCysAsnThrSerCysLeuPheGlnProArgGlyTYrGluCysArgAsp 67  
DB 205 GGGCCCTGCTGAACAATACCTCATGCTTTTTCAGCCACGAGGATATGAAGCCGGGAT 264

QY 68 AlaValAsnGluCysAspIleThrGluTyrCysThrGlyAspSerGlyGlnCysProPro 87  
DB 265 GCTGTGAACGAGTGCATATTTCTGAATTTTACTGAGAGCTCTGGTCAGTCCACCA 324

QY 88 AsnLeuHisLysGlnAspGlyTYrValCysAsnGlnAsnGlnGlyArgCys 104  
DB 325 AATCTTCATTAACGACGAGATATGATCATCAAAATCAGGCGGCTGC 375

RESULT 10  
ID AAZ45827 standard; DNA; 2088 BP.

AC AAZ45827;  
XX  
DT 25-APR-2000 (first entry)

DE Degenerate DNA encoding a disintegrin homologue designated zdint1.  
XX  
XX Human; disintegrin homologue; zdint1; cardiac myocyte; adipocyte;  
KW gene therapy; cell-cell interaction; chondro sarcoma; atherosclerosis;  
KW Alzheimer's disease; restenosis; ischemic reperfusion; obesity;  
KW intimal hyperplasia; tumour; platelet aggregation; apoptosis;  
KW neurogenesis; myogenesis; arthritis; myogenesis; neurogenesis;  
KW connective tissue disorder; chondrogenesis; tumour proliferation;  
KW inflammation; ss.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX WO200002912-A2.  
XX  
XX PD 20-JAN-2000.  
XX  
XX PF 09-JUL-1999; 99WO-US15638.  
XX  
XX PR 10-JUL-1998; 98US-0113883.  
XX  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX  
XX PI Sheppard PO, Baindur N, Deisher TA, Bishop PD;  
XX  
XX WPI; 2000-160898/14.  
XX  
XX Polypeptide useful in modulating cell-cell interaction in tissues of  
PT heart, brain, spinal cord and treating chondro sarcoma,  
PT atherosclerosis, restenosis, obesity, intimal hyperplasia and tumors -  
XX  
XX  
PS Disclosure; Page 129-130; 132pp; English.

CC The present sequence represents a degenerate sequence which encodes  
CC a human disintegrin homologue, designated zdint1. The zdint1 polypeptide  
CC is a cardiac myocyte proliferation and differentiation stimulator, as  
CC well as an adipocyte proliferation and differentiation inhibitor.  
CC Polynucleotides encoding zdint1 are used in gene therapy. The zdint1  
CC polypeptide is useful in modulating cell-cell interactions of cells  
CC derived from tissues of heart, brain, spinal cord and skeletal muscle.  
CC It is useful in treating and diagnosing chondro sarcoma, atherosclerosis,  
CC Alzheimer's disease, restenosis, ischemic reperfusion, obesity, intimal  
CC hyperplasia and tumors of heart, brain, and spinal cord. The zdint1

CC polypeptide is also useful in identifying its new family members,  
CC antagonists, agonists and antibodies. Antagonists, antibodies and fusion  
CC proteins of zdint1 are useful in inhibiting platelet aggregation,  
CC apoptosis, neurogenesis and myogenesis. Agonists and antagonists are  
CC useful in studying cell-cell interactions, arthritis, myogenesis,  
CC neurogenesis, connective tissue disorders, chondrogenesis, tumour  
CC proliferation and suppression, extracellular matrix proteins, repair  
CC and remodeling of ischemic reperfusion, inflammation, and apoptosis.

XX  
SQ Sequence 2088 BP: 412 A; 178 C; 381 G; 262 T; 855 other;

Alignment Scores:  
Pred. No.: 9, 81e-49 Length: 2088  
Score: 567.00 Matches: 91  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 13  
Query Match: 90.00% Indels: 0  
DB: 21 Gaps: 0

US-09-634-252a-4\_copy\_496\_599 (1-104) x AAZ45827 (1-2088)

QY 1 ProThrLysLeuPheGluProThrGluCysGlyAsnGlyTyrValGluAlaGluGlu 20  
DB 1117 CCNACNAAATYNTTGGACCCNACNGARTGCGNAAAGGNTAVGTNGARGCNGGARGAR 1176

QY 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysCysLysLysCysSerLeuSer 40  
DB 1177 TGXGATGCGGNTTYCAAGTNGARTGATGATGATGATGATGATGATGATGATGATGAT 1236

QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
DB 1237 AAYGNGCNCAYGTGWSNGAYGNCNCNTGYGAAYAAACNWSNTGYTNTTTCARCCN 1296

QY 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrCysThrGly 80  
DB 1297 MNGGNTAVGARTGTMGNGAVGCMGTNAAYGARTGAYATHAACGARTATTTGAACNGN 1356

QY 81 AspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrValCysAsnGlnAsn 100  
DB 1357 GAYWSNGNCARTGVCNCCNCAAYTNCAYAAARCARGAYGNTGYAAACARAY 1416

QY 101 GlnGlyArgCys 104  
DB 1417 CARGNGMGTGY 1428

Search completed: October 21, 2003, 10:34:11  
Job time : 278.138 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:21:31 ; Search time 2467.22 Seconds  
(without alignments)  
1024.499 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599  
Perfect score: 630  
Sequence: 1 PTKLFEPTEPCNGYVEAGEE.....CPNLHKODGYACNONOGRG 104

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/sgn2.1/USPTO/US09634252/runat\_21102003\_090324\_3381/app\_query.fasta\_1.462  
-DB=EST -QFMT=fastap -SUFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biom62 -TRANS=human10.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=tbl -NOM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBS=US09634252 @CGC 1.1 3596 @runat\_21102003\_090324\_3381 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: em\_estba: \*  
2: em\_estlum: \*  
3: em\_estlin: \*  
4: em\_estlm: \*  
5: em\_estcov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estlum: \*  
16: em\_estlum: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rod: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vrt: \*  
28: gb\_gss1: \*

29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	728	13	BUS39967
2	614	97.5	470	14	CB733332
3	611	97.0	2891	11	AK046677
4	611	97.0	3433	11	AK029301
5	611	97.0	3498	11	AK034022
6	578	91.7	340	14	F08148
7	563	89.4	722	12	B184544
8	560	88.9	752	10	BF691008
9	541.5	86.0	648	13	BUS36219
10	541.5	86.0	660	13	BUS39568
11	535.5	85.0	661	13	BUS38998
12	526.5	83.6	326	14	R15038
13	517.5	82.1	648	13	BUS31413
14	509	80.8	553	10	AW961362
15	502	79.7	405	14	R52569
16	492.5	78.2	664	12	B1429326
17	485	77.0	261	9	AA317222
18	480.5	76.3	652	12	B1429301
19	479.5	76.1	616	12	B1476560
20	465.5	73.9	598	12	B1981504
21	464	73.7	652	9	AV340595
22	458.5	72.8	599	12	BM070858
23	443.5	70.4	583	12	BM166739
24	438.5	69.6	579	12	B1844587
25	406.5	64.5	513	14	CB716849
26	404.5	64.2	443	14	CB788730
27	404.5	64.2	609	14	CB580769
28	404.5	64.2	704	14	BY722939
29	397	63.0	432	12	B1018895
30	387	61.4	400	14	CB523459
31	385.5	61.2	397	10	AM898896
32	384.5	61.0	618	12	BJ495352
33	378	60.0	402	12	B1019679
34	375	59.5	690	12	BJ059956
35	370.5	58.8	471	14	CB732873
36	344	54.6	428	9	AA718688
37	340.5	54.0	481	13	BY255403
38	338	53.7	552	9	AW660536
39	337	53.5	447	10	BF905823
40	317.5	50.4	622	14	CB578552
41	289.5	46.0	512	12	B1839501
42	287.5	45.6	570	10	BB611717
43	286.5	45.5	546	12	BP009939
44	283	44.9	570	13	BO833435
45	278.5	44.2	703	14	CB512918

#### ALIGNMENTS

RESULT 1  
BUS39967  
LOCUS BUS39967  
DEFINITION AGENCOURT 10254064 NIH\_MGC\_128 Homo sapiens cDNA clone  
IMAGE:6570971 5', mRNA sequence.  
ACCESSION BUS39967  
VERSION BUS39967.1 GI:22850408  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 728)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: NCI  
 CDNA Library Preparation: Michael Brownstein Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
<http://image.llnl.gov>  
 Place: LNCM2760 row: 1 column: 11  
 High quality sequence stop: 527.  
 Location/Qualifiers

## FEATURES

source

1..728  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6570971"  
 /issue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="DH10B (TI-phage-resistant)"  
 /clone\_1ib="NIH\_MGC\_128"  
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatagggc);  
 Site 2: SfiI (ggcgccctcgcc); Double-stranded cDNA was  
 prepared from a pool of 40 cell line polyA+ RNAs (bladder  
 - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
 kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
 ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
 gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
 used in cloning as follows:  
 5'-ATTCAGGCGGCGAGCGCGCGCATTCAGCGCGCG-3' and  
 5'-AAGCAGGCGGCGAGCGCGCGCGCATTCAGCGCGCG-3'. Full-length  
 enriched library was constructed using the Clontech  
 Creator SMART Kit and size-selected to contain the >2 kb  
 size fraction (other fractions present in NIH\_MGC\_126 and  
 NIH\_MGC\_127). Library created in the laboratory of T.  
 Uedlin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC  
 library"

BASE COUNT 213 a 142 c 210 g 162 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,476-54 Length: 728  
 Score: 630.00 Matches: 104  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x BUS39967 (1-728)

QY 1 ProThrlvsLeuPheGluProThrlGluCySGlYAsnglYTYrYalGluAglYglUglU 20  
 DB 257 CCACAAACACTATTAGCCACGAGAAATGTGAATGATACCTGGAACCTGGGAGAG 316  
 QY 21 CysAspCySGlYPhelIsvAlGluCyTYrGlyLeuCyCySlyslsCySserLeuSer 40  
 DB 317 TGTGATTGGTTTATGATGTGAATCTATGATTAAGCTGTAAGAAAGTTCCCTCC 376  
 QY 41 AsnglYAlAhIsCySserAspGlyProCyCySAsnAsnThsSerCySleuPheGlnPro 60  
 DB 377 AACGGGGCTCAGCGACGACGCGGCGCTGCTGTAAACATACCTCACTCTTTTCGCCA 436  
 QY 61 ArgGlyTYrGluCySArGAspAlaValAsnglYCyAspIleThrGluTYrCySthrgly 80  
 DB 437 CGAGGCTATGAATGCGGGATGCTGTGAACGAGTGTGATATTAAGTAATTTGATCGGA 496  
 QY 81 AspSerGlyGlnCySProPheAsnLeuHIslySglAspGlyTYrYalCySAsnglnAsn 100  
 DB 497 GACTCTGTGATGAGCCACCAAAATCTTCAATAGCAAGAGGATGCAATGCAATCAAAAT 556

QY 101 GlnGlyArgCys 104  
 DB 557 CAGGGCCCGCTGC 568

RESULT 2  
 LOCUS CB733332  
 DEFINITION AMGNNUC:NRH5-00422-H8-A W Rat hypothalamus (10471) Rattus  
 norvegicus cDNA clone nrh5-00422-h8 5', mRNA sequence.  
 ACCESSION CB733332  
 VERSION CB733332.1 GI:29800497  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE  
 1 (bases 1 to 470)  
 AUTHORS Amgen EST Program.  
 TITLE Amgen Rat EST Program  
 JOURNAL Unpublished  
 COMMENT Contact: Dan Fitzpatrick  
 Amgen, Inc  
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
 Tel: 805 447-4881  
 Plate: 00422 row: h column: 8.

## FEATURES

source

1..470  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="nrh5-00422-h8"  
 /note="Vector: pSPORFL; Site 1: SalI; Site 2: NotI; W Rat  
 hypothalamus adult female Wistar rat avg. Insert size 2.3  
 kb fraction 6 and 7"

BASE COUNT 124 a 111 c 137 g 98 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,766-52 Length: 470  
 Score: 614.00 Matches: 101  
 Percent Similarity: 98.08% Conservative: 1  
 Best Local Similarity: 97.12% Mismatches: 2  
 Query Match: 97.46% Indels: 0  
 DB: 14 Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x CB733332 (1-470)

QY 1 ProThrlvsLeuPheGluProThrlGluCySGlYAsnglYTYrYalGluAglYglUglU 20  
 DB 57 CCGACTAAGCGCTTTGAGCCACGCAATGTGGAACGATATGTGGAACCGGAGAGAA 116  
 QY 21 CysAspCySGlYPhelIsvAlGluCyTYrGlyLeuCyCySlyslsCySserLeuSer 40  
 DB 117 TCGATTGTGGTTTCCAGTGTGAATGCTAAGCGAGATTCCTGTAAAGAAATGCTCTTCC 176  
 QY 41 AsnglYAlAhIsCySserAspGlyProCyCySAsnAsnThsSerCySleuPheGlnPro 60  
 DB 177 AACGGGGCCACATGAGAGGCGCCCTGTGTAAACAACCTGCTGCTTTTCAGTCA 236  
 QY 61 ArgGlyTYrGluCySArGAspAlaValAsnglYCyAspIleThrGluTYrCySthrgly 80  
 DB 237 CGAGGCTACGAATGCGGGATGCCCTGAACAGCTGTATTAACGAGTACTGACCCGA 296  
 QY 81 AspSerGlyGlnCySProPheAsnLeuHIslySglAspGlyTYrYalCySAsnglnAsn 100  
 DB 297 GACTCTGCGACGATGCGCCCGCAATCTCATTAACAAGATGCTAGCGCTGCAATCAAAAT 356  
 QY 101 GlnGlyArgCys 104  
 DB 357 CAGGGTCCGCTGC 368

RESULT 3	AK046677	2891 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	AK046677				
DEFINITION	Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length enriched library, clone:B43030710 product:a disintegrin and metalloproteinase domain 23, full insert sequence.				
ACCESSION	AK046677				
VERSION	AK046677.1	GI:26091652			
KEYWORDS	HTC, CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Teshiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arkawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, Y., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Katukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pongle, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Butt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombereis, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Teyo-Oka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilming, L., Wyshew-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsaki, S. and Hayashizaki, Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085660				
PUBMED	11217851				
REFERENCE	5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
MEDLINE	6 (bases 1 to 2891)				
PUBMED	Adachi, J., Aizawa, K., Akimura, T., Arkawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hasegawa, Y., Hasegawa, M.,				

	TITLE	Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukabe, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Niishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Ozaki, Y., Saico, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, T., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.					
	JOURNAL	Submitted (16-JULY-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail: genome-res@gs.c.riken.go.jp], URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)					
	COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://tancom.gsc.riken.go.jp/. Location/Qualifiers					
FEATURES	source	1..2891 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:B430307120" /db_xref="Taxon:10090" /clone="B430307120" /sex="male" /tissue_type="adipose" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="4 days neonate" 1..2891 /note="A disintegrin and metalloprotease domain 23 (MG MG1:1345162, GB NM_011780, evidence: BLASTN, 100%, match=2431)"					
	BASE COUNT	721 a	794 c	778 g	598 t		
	ORIGIN						
	Alignment Scores:	Pred. No.: 3.74e-51	Length: 2891				
	Score:	611.00	Matches: 100				
	Percent Similarity:	98.08%	Conservative: 2				
	Best Local Similarity:	96.15%	Mammals: 2				
	Query Match:	96.98%	Indels: 0				
	DB:	11	Gaps: 0				
	US-09-634-252A-4_COPY_496_599 (1-104) x AKO46677 (1-2891)						
Oy	1	ProThrIysLeuPheGluIProThrcIuCySgIyaSngIyTYrValGIuAlaGIyGIuGIu	20				
Dd	1798	CCAACTAAGCTGTGGTAGCCCGGAATGTGGAATAATGCATATGTGGAGCGCGGAGGA	185				
Oy	21	CysAspCYSGIYPheHisValaIGIuCYSTyrGLyLeuCYGSylvLYlcySSetleusEr	40				
Dd	1858	TGCGACTGTGTTTCCATGTGGAATGCATATGGAAGTTTCCTGTAAAGATGTTCCGCTTCC	191				
Oy	41	AsnGIyAlaHisCYSSerASpGIYProcCYsCysAsnaEnthrSerCYsLeuPheGlnPro	60				
Dd	1918	AATGGGCCCACTGCATGACGCGCCCTCGCTGTAAACAACACTCATATGCTTTTATAGTCA	197				
Oy	61	ArgGIyTYrGIuCYsarGAsPaLaValasngIuCYsaSpIIerthGIuTYrCYstHsily	80				
Dd	1978	CGAGGGTATGATATGCGGATGCGGTAAACAGCTGTGATTATCACCGAATGACACTGGA	203				
Oy	81	AsperGIyGIuCYePrOProAsnLeuHisLySGInaSPGIYTYrAlaCYaAsngInaSn	100				

Db 2038 GACTGTGGCAGTCCACCGAAGCTCCATAAACAAGATGCTATAGTCGAATCAAAAT 2097  
QY 101 GINGLYARGCYs 104  
Db 2098 CAGGCTCGCTGC 2109

RESULT 4  
AK029301  
LOCUS  
DEFINITION AK029301 3433 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched  
library, clone:432416K07 product: a disintegrin and metalloprotease  
domain 23, full insert sequence.

ACCESSION  
VERSION AK029301.1 GI:26081272  
KEYWORDS AK029301  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE Carninci, P. and Hayashizaki, Y.  
JOURNAL High-efficiency full-length cDNA cloning  
MEDLINE Meth. Enzymol. 303, 19-44 (1999)  
PUBMED 99279253  
REFERENCE 10349636

TITLE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
JOURNAL Normalization and subraction of cap-trapper-selected cDNAs to  
MEDLINE Prepare full-length cDNA libraries for rapid discovery of new genes  
PUBMED Genome Res. 10 (10), 1617-1630 (2000)  
REFERENCE 20499374  
AUTHORS 11042159

TITLE 3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kutsuna, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
JOURNAL RIKEN integrated sequence analysis (RISA) system-384-Format  
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 1076661

TITLE 4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batelov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staudli, F., Furuno, M., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, P., Mombaerts, P., Nordone, P.,  
Rising, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.  
and Hayashizaki, Y.  
JOURNAL Functional annotation of a full-length mouse cDNA collection  
MEDLINE Nature 409 (6821), 685-690 (2001)  
PUBMED 11217851  
REFERENCE 5  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation  
JOURNAL of 60,770 full-length cDNAs  
REFERENCE Nature 420, 563-573 (2002)  
AUTHORS 6 (bases 1 to 3433)

Adachi, J., Aizawa, K., Akimura, T., Arai, A., Hasegawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirokane, T.,  
Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Konda, M.,  
Koyra, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission  
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
COMMENT cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
FEATURES  
source 1..3433  
location/Qualifiers  
1..3433  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM:DB:4832416K07"  
/db\_xref="taxon:10090"  
/clone="4832416K07"  
/cissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
misc\_feature 1..3433  
/note="a disintegrin and metalloprotease domain 23  
(MSD|MG:1345162, GBLNM\_011780, evidence: BLASTN, 100%,  
match=2431)"

BASE COUNT 944 a 792 c 836 g 861 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.68e-51 Length: 3433  
Score: 611.00 Matches: 100  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 96.15% Mismatches: 2  
Query Match: 96.98% Indels: 0  
DB: 11 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AK029301 (1-3433)

QY 1 ProthrluysleuphegluProthrluysleuysnglytyrvalglualglglu 20  
Db 1099 CCACTTAAGCTGTTTGAACCCACGAAATGTGAATGATATGTGAGAGCGGGAGGAA 1158

QY 21 CysaspCygslyPnehisvalglucyTytylgluCyCylyslVysCysSerleuser 40  
Db 1159 TGCAGCTGTGTTTCCATGCGAATGCTATGAGACTTTGCTTAAGAACTGTTCCCTCC 1218

QY 41 AsnglylaHsCysSerAspGlyProCyCySasnsantrrsCysleuphegluPro 60  
Db 1219 AATGGGCGCCACTGAGGAGCGCCCTGCTGTAAACAACACCTCATGCTTTTCAGTCA 1278

QY 61 ArgglyTyrgluCysArgAspAlaValasngluCysAspIleThrGluTytyrCysThrGly 80

Db 1279 CGAGGATGATGATTCGGAGATCCGTAACAGCTGTGATATACACCGATGACGACGGA 1338

Qy 81 Abspserglyncysproproanleuhishysginspglytcalacysasn 100

Db 1339 GACCTGCGCGATGCTCCACGACCACTCCATTAACAGATGCTAGCTGACATCAAAAT 1398

Qy 101 GlnGlyAArgCys 104

Db 1399 CAGGCTCGCTGC 1410

RESULT 5

AK034022

LOCUS

DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330142F22 product: a disintegrin and metalloprotease domain 23, full insert sequence.

ACCESSION AK034022

VERSION AK034022.1 GI:26329622

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Ichih, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanaka, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

TITLE JOURNAL MEDLINE PUBMED PUBLISHED 20530913 11076861

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, T., Peesle, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagnier, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, N., Baldarelli, R., Baren, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., But, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, C., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Mochizuki, L., Washima, J., Mazza, J., Mombauris, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, T., Sakamoto, N., Sasaki, H., Sato, K., Schenbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshew-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5

AUTHORS

TITLE

JOURNAL

REFERENCE

ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W., HAYASHIDA, K., HAYATSU, N., HITAMOTO, K., HIRAKO, T., HIROZANE, T., HORII, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOUDE, M., KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHNATO, N., OKAZAKI, Y., SAITO, R., SAITOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SAKAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKIHARA, S., TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., MURAMATSU, M. and HAYASHIZAKI, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshinide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1..3498

/organism="Mus musculus"

/mol\_type="mRNA"

/script="C57BL/6J"

/db\_xref="FANTOM\_DB:9330142F22"

/db\_xref="taxon:10090"

/clone="9330142F22"

/sex="male"

/tissue="dienecephalon"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

303..2375

/note="unannoted protein product: a disintegrin and metalloprotease domain 23 (MGI|MI:1345162, GBNM\_011780, evidence: BLASTN, 100%, match=2431)

putative"

/protein\_id="BAC28550.1"

/codon\_start=1

/db\_xref="GI:26329623"

/translation="MKPRTGISRRSLTSGSCTGCGRCRSPGPPAPAPRCRLTLLVLLPALATSSRRPARGAAPSAPHNENETKTLGVLADENLTLOQSSRTSYSAVKEKTLTPSLRVLYINODESPHYVLDYARQQRNKVLAQAFQLEAFGSKFLIDTLNUNGSLSSDYVEVLYHEDGKQMSYSGEHYKISGVDSRVSLSTCQNLHGMFEEDTFVYVMEPELTDEKSTGRPHIIQTLGQYSKQKNIJSTDSDDMLPEIQMLRRKRAVNSPGLVEPEMEYLEMIYNDKRTYKRKSSHAHTNNFASVNLVDSIYKEDLNRYVAVVETWKEKHIDITINPNVMDPSFYRORICOHDAVHLISVTFH YKSSLSYFEGVCSRRIGCVNVEYGLPMAYQVLSGLAONLGIQMEPPSRKPCCEI ESNGKSMETGTGVSRSKFSKCSILEFRDLQGGGACLPNRPKLFEPEPCGNGVE AGEBCDGGFVEPEVGCCKKCSLNGHAGDCNNNTSCLFQSRGECDAVNSCDI TEYCTGSGCCPNLHKQDGYSCNQGRCYVNECEKTRDQCYIWTAKAGDSKFCY EKXNTEETKGNCGKXKQDRIWIPCSKIDVFCGFLCTNLTPARPIGQLGSEIIPSPFVH QGVADR"

BASE COUNT 906 a 880 c 853 g 859 t

ORIGIN

## Alignment Scores:

Pred. No.: 4,79e-51 Length: 3498  
Score: 611.00 Matches: 100  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 96.15% Mismatches: 2  
Query Match: 96.98% Indels: 0  
DB: 11 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x AK034022 (1-3498)

QY 1 ProthylveleuphgluProthrgluCysGlyAsnGlyTyrValGluAlaGluGluGlu 20  
DB 1779 CCAACTAAGCTGTTGAGCCCAAGGATGTGAAAATGATGTGAGGCGGGAGGNA 1838  
QY 21 CysAspCysGlyPheHisValGluCysTyrGlyLeuGlyLeuGlySerLeuSer 40  
DB 1839 TCCGACTGCTGTTCCATGTGAAATGCTTGAAGTTGCTGTAAGAAGCTTCCTCC 1898  
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
DB 1899 AATGGGGCCCACTGCAGTACGCGCCCTGCTGTAAACACCTCATGCTTTTTCAGTCA 1958  
QY 61 ArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrCysThrGly 80  
DB 1959 CGAGGCTATGAATGTGGGATGCCCTAAACAGCTGTATACCCAGTACTGCTGGA 2018  
QY 81 AspSerGlyGlnCysProProAsnLeuHisGlyGlnAspGlyTyrAlaCysAsnGlnAsn 100  
DB 2019 GACTCTGGCCAGTCCCACTGCAACCTCCATTAAACAGATGGCTATAGCTTCATCAAAAT 2078  
QY 101 GlnGlyArgCys 104  
DB 2079 CAGGCTCGCTGC 2090  
RESULT 6  
F08148  
LOCUS F08148 340 bp mRNA linear EST 21-FEB-1995  
DEFINITION HSCRF011 normalized infant brain cDNA Homo sapiens cDNA clone  
c-2rf01, mRNA sequence.  
ACCESSION F08148  
VERSION F08148.1 GI:677656  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
REFERENCE 1 (bases 1 to 340)  
AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houligatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y., Sebastiani-Kabakchie, C. and Tessier, A.  
TITLE IMAGE: molecular integration of the analysis of the human genome and its expression  
JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
MEDLINE 95277534  
PUBMED 7757816  
COMMENT Contact: Genethon  
Genethon Centre de recherche sur le Genome Humain  
1, rue de l'Internationale, BP60 91002 Evry Cedex, FRANCE  
Tel: 33169472800  
Fax: 33160778698  
Email: genexpress@genethon.fr  
Single read.  
Genexpress\_library\_id: C; Genexpress\_sequence\_id: y1c-2rf01  
Seq primer: (-21)M13-universal.  
Location/Qualifiers  
1..340  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="C-2rf01"  
/sex="Female"

/tissue\_type="total brain"  
/dev\_stage="3 months old"  
/clone\_lib="normalized infant brain cDNA"  
/note="Organ: brain; Vector: lambda B; Site\_1: HindIII;  
Site\_2: NotI; sex=Female; dev\_stage=3 months old;  
Isolate=muscular atrophy patient; tissue\_type=total brain  
; total mRNA was oligo-(dT) primed and directionally  
cloned 5' -> 3' into the HindIII -> NotI sites of the  
lambda B vector. Clone library from B.Scores, Psychiatry  
Dept. Columbia University, USA  
Bento Scores, P.N.A.S in press"  
BASE COUNT 87 a 70 c 90 g 84 t 9 others  
ORIGIN

## Alignment Scores:

Pred. No.: 5,09e-49 Length: 340  
Score: 578.00 Matches: 97  
Percent Similarity: 94.17% Conservative: 0  
Best Local Similarity: 94.17% Mismatches: 6  
Query Match: 91.75% Indels: 0  
DB: 14 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x F08148 (1-340)

QY 2 ThrIysLeuphgluProthrgluCysGlyAsnGlyTyrValGluAlaGluGluGluCys 21  
DB 27 ACAAGCTATTATTTNAGCCACCGAATGTGAATGNTMCGTGAAGCTGGGAGAGTGT 86  
QY 22 AspCysGlyPheHisValGluCysTyrGlyLeuGlyCysGlySerLeuSerLeu 41  
DB 87 NATGTGTTTATGATGGAATGCTATGATATTATCTAAGAAATGTCCTCCCAAC 146  
QY 42 GlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnProArg 61  
DB 147 GGGGCTCACTGACGAGGAGCGGCGCTGCTGTAACTACTGCTTTTTCAGCCACA 206  
QY 62 GlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrCysThrGlyAsp 81  
DB 207 GGGTATGAATGCGCGGATGCTGTGAACGAGTGTATTAAGTAATATTGACTGAGAC 266  
QY 82 SerGlyGlnCysProProAsnLeuHisGlyGlnAspGlyTyrAlaCysAsnGlnAsn 101  
DB 267 TCTGTGATGAGTCCCACTCAATCTTCATTAAGACAGATATGATGCAATCAAAATCAG 326  
QY 102 GlyArgCys 104  
DB 327 GCGCCCTGC 335  
RESULT 7  
B1184544  
LOCUS B1184544 722 bp mRNA linear EST 10-JUL-2001  
DEFINITION UNT-P-FN-co-e-05-0-UNT\_61 UNT-P-FN Sus scrofa cDNA clone  
UNT-P-FN-co-e-05-0-UNT\_3', mRNA sequence.  
ACCESSION B1184544  
VERSION B1184544.1 GI:14658953  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
REFERENCE 1 (bases 1 to 722)  
AUTHORS Caetano, A.R., Johnson, R.K. and Pomp, D.  
TITLE Generation and sequence characterization of a normalized cDNA  
library from swine ovarian follicles  
JOURNAL Mamm. Genome 14 (1), 65-70 (2003)  
MEDLINE 22419904  
PUBMED 12532269  
COMMENT Contact: Pomp, D  
Department of Animal Science  
University of Nebraska, Lincoln  
Lincoln, NE 68583-0908, USA  
Tel: 402 472 6416  
Fax: 402 472 6362

Email: dpoomp@unl.edu  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The following repetitive elements were found in this cDNA sequence: 3-43, >(TAAA)n#Simple\_repeat  
Seq primer: M13 -29  
POLYA=yes.

## FEATURES

source

## Location/Qualifiers

```
1..722
/organism="Sus scrofa"
/mol_type="mRNA"
/strain="University of Nebraska, Lincoln Swine Selection
lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-co-e-05-0-UNL"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UNL-P-FN"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genoms Research 6: 791-806
, 1996.
TAG_SEQ=None found"
```

BASE COUNT 188 a 181 c 131 g 222 t

ORIGIN

## Alignment Scores:

Pred. No.: 4.46e-47 Length: 722  
Score: 563.00 Matches: 92  
Percent Similarity: 97.89% Conservative: 1  
Best Local Similarity: 96.84% Mismatches: 2  
Query Match: 89.37% Indels: 0  
DB: 12 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x B1184544 (1-722)

```
OY 10 CysGlyAsnGlyTyrValGluAlaGlyGluGluCysAspGlyPheHisValGluCys 29
Db 721 TGTGAAACGGCTTATCTGAGAGCTGGAGGAGATGCGATGGGGATCAGTGTGAAATGC 662
OY 30 TyrGlyLeuCysCysLeuSerLeuSerAsnGlyAlaHisCysSerAspGlyPro 49
Db 661 TACGACACTGTGTCAGAAATGCTCTCTCTCAACGGATGCCCATTTGACAGCGAGGCC 602
OY 50 CysCysAsnAsnThrSerCysLeuPheGlnProArgGlyTyrGlyCysArgAspAlaVal 69
Db 601 TGCCTTAATAGTACTCTGCTCTTTTTCAGCCACCTGGGTATGATGTCGGGATCTCTGTG 542
OY 70 AsnGlyCysAspGlyLeuThrGlyTyrCysThrGlyAspSerGlyGlnCysProProAsnLeu 89
Db 541 AATGATATGATATTAACCGAATATTTACTGAGAGCTCTGGCCACATGCCCAATCTT 482
OY 90 HisLysGlnAspGlyTyrAlaCysAsnGlnAsnGlnGlyArgCys 104
Db 481 CATAGCAAGATGTTATGATGCAATCAAAATCAGGCCCGCTGC 437
```

RESULT 8  
BF691008  
LOCUS 602247043F1 NIH\_MGC\_62 Homo sapiens cDNA clone IMAGE:4332275 5',  
DEFINITION mRNA sequence.  
ACCESSION BF691008  
VERSION BF691008.1 GI:11976416  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 752)  
AUTHORS  
NH-MGC <http://mgc.nci.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC/DCPD/DRP  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.jnl.gov>  
Plate: LNCM198 row: e column: 12  
High quality sequence stop: 719.

## FEATURES

source

## Location/Qualifiers

```
1..752
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4332275"
/issue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (TI phase-resistant)"
/clone_lib="NIH MGC 62"
/notes="Organ: skin; Vector: pNNR-LIB (Clontech); Site 1:  
SfiI (ggccgctcggcc); Site 2: SfiI (ggccattatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCATTTAGGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGGAGCGGCCGACATG-dt(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."
```

BASE COUNT 195 a 151 c 217 g 189 t

ORIGIN

## Alignment Scores:

Pred. No.: 9.46e-47 Length: 752  
Score: 560.00 Matches: 92  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 88.89% Indels: 0  
DB: 10 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x BF691008 (1-752)

```
OY 13 GlyTyrValGluAlaGlyGluGluCysAspGlyPheHisValGluCysTyrGlyLeu 32
Db 1 GGATACGTGGAAGCTGGGAGAGTGTGATTTGTTTATGTAATGTAATGATTA 60
OY 33 CysCysLysCysCysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsn 52
Db 61 TGCCTTAAGAAATTTCCCTCTCCACAGGGCTACTGACGACGAGGCCCTGCTGTAC 120
OY 53 AsnThrSerCysLeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCys 72
Db 121 AATACCTCATGCTCTTTTTCAGCCACGAGGATGATATGCGGATGCTGTACAGAGTGT 180
OY 73 AspIleThrGlyTyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLysGln 92
Db 181 GATATTACTGAATATTTACTGAGAGCTGTGTAGTCCACAAATCTTTCATTAACCA 240
OY 93 AspGlyTyrAlaCysAsnGlnAsnGlnGlyArgCys 104
Db 241 GACGATATGATCATCAATAATCAGGCCCGCTGC 276
```

RESULT 9  
BU396219



LOCUS BU396219 648 bp mRNA linear EST 27-NOV-2002  
 DEFINITION 603805442F1 CSEQCHN57 Gallus gallus cDNA clone CHEST783d12 5', mRNA  
 ACCESSION BU396219  
 VERSION BU396219.1 GI:25765275  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 648)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken CDNAS  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 )  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1..648  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST783d12"  
 /dev\_stage="16 day embryo"  
 /lab\_host="DH10B"  
 /clone\_11b="CSEQCHN57"  
 /note="Organ: Brain; Vector: pBluescript II KS(+); Site\_1:  
 EcoRI; Site\_2: NotI; This normalized library was  
 constructed from 1 million independent clones. cDNA  
 synthesis was initiated using an oligo(dT) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunt-ended, ligated to NotI adapters, digested with EcoRI  
 , size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."  
 BASE COUNT 187 a 117 c 169 g 175 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,81e-45 Length: 648  
 Score: 541.50 Matches: 91  
 Percent Similarity: 91.35% Conservative: 8  
 Best Local Similarity: 87.50% Mismatches: 8  
 Query Match: 85.95% Indels: 1  
 DB: 13 Gaps: 1  
 US-09-634-252a-4\_COPY\_496\_599 (1-104) x BU396219 (1-648)  
 QY 1 ProThrLysLeuPheGluProThrGluCysGlyAsnGlyTyrValGlnAlaGluGluGlu 20  
 DB 228 CCAACAAGCTCTTCAAACTAGTGTGGAATGATGATGTGACAAACAGAGAGAAGA 287  
 QY 21 CysAspCysGlyPheHisValGluCysTyrGlyLeuCysCysLysLysCysSerLeuSer 40  
 DB 288 TCGCATGGCGTTTCCGATGGAATGCTACGACAGACTGTGTGAAGAAGTCTCTTTT 347  
 QY 41 AanglyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60

DB 348 AATGAGCTCACTGATGATGAGCCCTGCTGT---AATAGCTGCTGCTTTTTTTCGA 404  
 QY 61 ArgGlyTyrGluCysArgAlaValAsnGluCysAspIleThrGluTyrCysThrGly 80  
 DB 405 CGAGCTTGAAGTCTGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 464  
 QY 81 AspSerGlyGlnCysProPheAsnLeuHisLysGlnAspGlyTyrAlaCysAsnGlnAsn 100  
 DB 465 GATTTCGCCAGATGTCACCAATATCTTCATTAACAAGATGATATGCTTGATTTCTAAT 524  
 QY 101 GInGlyArgCys 104  
 DB 525 CAGGAGCGCTGC 536  
 RESULT 10  
 LOCUS BU395688 660 bp mRNA linear EST 27-NOV-2002  
 DEFINITION 603805095F1 CSEQCHN57 Gallus gallus cDNA clone CHEST782n24 5', mRNA  
 ACCESSION BU395688  
 VERSION BU395688.1 GI:25764744  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 660)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken CDNAS  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 )  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1..660  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST782n24"  
 /dev\_stage="16 day embryo"  
 /lab\_host="DH10B"  
 /clone\_11b="CSEQCHN57"  
 /note="Organ: brain; Vector: pBluescript II KS(+); Site\_1:  
 EcoRI; Site\_2: NotI; This normalized library was  
 constructed from 1 million independent clones. cDNA  
 synthesis was initiated using an oligo(dT) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunt-ended, ligated to NotI adapters, digested with EcoRI  
 , size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."  
 BASE COUNT 193 a 122 c 170 g 174 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.95e-45 Length: 660  
 Score: 541.50 Matches: 91  
 Percent Similarity: 91.35% Conservative: 4



Best Local Similarity: 87.50%  
 Query Match: 85.95%  
 DB: 13  
 Mismatches: 8  
 Indels: 1  
 Gaps: 1

US-09-634-252A-4\_COPY\_496\_599 (1-104) X BU395688 (1-660)

QY	1	ProthrllysleuPheGluProThrGluCysGlyAsnGlyTyrValGluAlaGlyGluGlu	20
DB	241	CCAAACAAGCTCTTGAACACTACTGAGTGGAAATGATATGTAGAAGCAGAGAGAA	300
QY	21	CysAspCysGlyPheHisValGluCysTyrGlyLeuCysCysLysLysCysSerLeuSer	40
DB	301	TGCGATTGCGGTTCCGAATGGAAATGCTACCGACAGCTGTGTAAAAAGTGCTCTTTCT	360
QY	41	AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro	60
DB	361	AATGAGCTCTCACTGTAGTGAATGGGCTTGCTGT--AATAGTGTGTCTTTTTCCTCA	417
QY	61	ArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrCysThrGly	80
DB	418	CGAGCTATGACTGTAGTATGACAGTGAATGTGTATTGCAAGTTCTGCACCTGGA	477
QY	81	AspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTyrAlaCysAsnGlnAsn	100
DB	478	GATTCTGGCCAGTGTCCACCAATCTTCATTAACAAGATGATAGCTTGTGATTCTTAAT	537
QY	101	GlnGlyArgCys	104
DB	538	CAGGAGCGTTGC	549

Search completed: October 21, 2003, 13:06:54  
 Job time : 2478.22 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:12:31 ; Search time 1815.35 Seconds  
(without alignment)  
1239.449 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGAHCSDGPCNNTSC.....CDIREYCTGDSGCPNLAHK 55

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xlh  
-Q=/cgm\_1/USPRO.spool/US09634252/runat\_21102003\_090323\_3370/app\_query.fasta\_1.462  
-DB=genemb1 -QWt=faetap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCLALIGN=200 -THR SCORE=PCC -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTMT=PCC -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09634252.QCGN.1.1.3608 @runat\_21102003\_090323\_3370 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pac:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcg\_hum:\*  
40: em\_hcg\_mus:\*  
41: em\_hcg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	1668	6 AX235014	AX235014 Sequence
2	332	100.0	2499	6 BD130423	BD130423 DNAs and
3	332	100.0	3054	6 E26532	E26532 Novel prote
4	332	100.0	3054	9 AB009672	AB009672 Homo sapi
5	332	100.0	3079	9 HSA5580	AJ005580 Homo sapi
6	332	100.0	4043	6 AX299710	AX299710 Sequence
7	319	96.1	2891	10 AB009673	AB009673 Mus muscu
8	213	64.2	1464	6 E10490	E10490 CDNA encodi
9	213	64.2	1464	6 E125845	E125845 Sequence 5
10	213	64.2	1464	6 E13482	E13482 Sequence 5
11	213	64.2	2908	6 H10492	H10492 Homo sapien
12	213	64.2	2913	6 E10492	E10492 CDNA encodi
13	213	64.2	2913	6 E125847	E125847 Sequence 7
14	213	64.2	2913	6 E13484	E13484 Sequence 7
15	213	64.2	2923	6 E10491	E10491 CDNA encodi
16	213	64.2	2923	6 E125846	E125846 Sequence 6
17	213	64.2	2923	6 E13483	E13483 Sequence 6
18	213	64.2	3168	9 AB009675	AB009675 Homo sapi
19	213	64.2	3183	6 E10493	E10493 CDNA encodi
20	213	64.2	3183	6 E125848	E125848 Sequence 8
21	213	64.2	3183	6 E13485	E13485 Sequence 8
22	213	64.2	3229	10 AB009676	AB009676 Mus muscu
23	185	55.7	1588	5 AF032384	AF032384 Xenopus 1
24	185	55.7	1674	6 AX235012	AX235012 Sequence
25	185	55.7	2604	6 BD130424	BD130424 DNAs and
26	185	55.7	2697	6 E26531	E26531 Novel prote
27	185	55.7	2786	9 AB009671	AB009671 Homo sapi
28	185	55.7	2796	9 AF155381	AF155381 Homo sapi
29	185	55.7	2805	6 E26530	E26530 Novel prote
30	185	55.7	2858	9 AF155382	AF155382 Homo sapi
31	185	55.7	3259	9 AF158637	AF158637 Homo sapi
32	185	55.7	3295	9 AF073291	AF073291 Homo sapi
33	184	55.4	2773	10 AB009674	AB009674 Mus muscu
34	183	55.1	3447	5 AF032383	AF032383 Xenopus 1
35	173.5	52.3	2444	10 RNTMDCVI	Y11492 R.norvegicu
36	173.5	52.3	2653	6 A51795	A51795 Sequence 30
37	173.5	52.3	2815	10 RNO521198	AJ251198 Rattus no
38	168.5	50.8	276	6 AX698112	AX698112 Sequence
39	168.5	50.8	1638	6 AX235004	AX235004 Sequence
40	168.5	50.8	1901	6 AK075498	AK075498 Homo sapi
41	168.5	50.8	2720	9 HSA46005	U46005 Human MDC15
42	168.5	50.8	2740	6 AX410713	AX410713 Sequence
43	168.5	50.8	2740	6 HSA41767	U41767 Human metlar
44	168.5	50.8	2819	9 BC014566	BC014566 Homo sapi
45	168.5	50.8	2968	6 AX565631	AX565631 Sequence

RESULT 1

ALIGNMENTS

AX235014  
LOCUS AX235014 1668 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 17 from Patent WO0162905.  
ACCESSION AX235014  
VERSION AX235014.1 GI:15593673  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Fanelow, W.C., Cerretti, D.P., Poindexter, K.M. and Black, R.A.  
Integrin antagonists  
Patent: WO 0162905-A 17 30-AUG-2001;  
IMMUNEX CORPORATION (US)  
LOCATION/Qualifiers  
1. 1668  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="fusion polypeptide"  
25. 1647  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC69803.1"  
/db\_xref="GI:15593674"  
/translation="METDTLLMVLVLLMWPSTGTSTGCGNGVEAGDCCGPHVCEY  
LCCKKCSLNGACSDGPCNNSTCLFOPRGYECDAINECDITPCYSGDCCPRL  
HKODYACNONGRCYNGECRARDNOCYITWTKAGSKDFEYELNTEBTKNGCGK  
DGDRIQCSKHDVFCGFLCTNLTRAPRIGQLQGIIPSEFHQGRVIDCSGAHVLL  
DDTDVYVEDGTGCGSMWCLDRKCIQIALMSSCPIDSKKVCSGHGVSNATCI  
CDFWAGDICSIRDPVRLHPKDEGPKSPKSTNSCDKHTCPCPAEEAGASVF  
LEPPKDTLMTSRTPEVTCVVDVSHDEPKFVWYDGVSVHNAKTRPEQVNST  
YRVSVLTVLHODMTNGKCYKCKVKNKALPAIETCTISKAGCPPOVYITPERSDE  
LTKNVSFLCLVKGFPSPDIAYEMESNQPENNYKTTTPVLDSDSFFLYSLKLTVDKS  
RMQGNVFSVMEALHMHYQKSLSPGK"  
BASE COUNT 423 a 444 c 455 g 346 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,18e-31 Length: 1668  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x AX235014 (1-1668)  
QY 1 CysSerLeuSerArgGlyAlaHisCysSerArgGlyProCysCysAsnAsnThrSerCys 20  
DB 172 TGTTCCTCTCAACGGGGCTCACTGCAGCAGCGGGCCCTGCTGTAACAATCCTCATGT 231  
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
DB 232 CTTTTCAGCCACGAGGGTATGAATGCCGGAGTGTCTGTAAAGAGTGTATTAATCTGAA 291  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
DB 292 TATTGTACTGAGACTCTGGTCAGTGCCACCAAAATCTTCATTAAG 336  
RESULT 2  
BD130423  
LOCUS BD130423 2499 bp DNA linear PAT 18-SEP-2002  
DEFINITION DNAs and polypeptides of metalloprotease disintegrins SVPH3-13 and  
SVPH3-17.  
ACCESSION BD130423  
VERSION BD130423.1 GI:22225368  
KEYWORDS JP 2002503472-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2499)  
AUTHORS Cerretti, D.P.  
TITLE DNAs and polypeptides of metalloprotease disintegrins SVPH3-13 and  
SVPH3-17  
JOURNAL Patent: JP 2002503472-A 2 05-FEB-2002;  
IMMUNEX CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002503472-A/2  
PD 05-FEB-2002  
PF 11-FEB-1999 JP 2000531569  
PI 11-FEB-1998 US 60/074310  
PC C12N15/09, C07K16/40, C12N1/19, C12N1/21, C12N5/10, C12N9/64 PC  
, C12O1/37, G01N33/68  
CC C12N15/00, C12N5/00  
CDS DNAs and polypeptides of metalloprotease disintegrins CC  
SVPH3-13 and SVPH3-17  
FH Key Location/Qualifiers  
FT source 1. 2499  
FT /organism="Homo sapiens (human)".  
FEATURES  
source  
1. 2499  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 677 a 593 c 645 g 584 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9,17e-31 Length: 2499  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x BD130423 (1-2499)  
QY 1 CysSerLeuSerArgGlyAlaHisCysSerArgGlyProCysCysAsnAsnThrSerCys 20  
DB 1594 TGTTCCTCTCAACGGGGCTCACTGCAGCAGCGGGCCCTGCTGTAACAATCCTCATGT 1653  
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
DB 1654 CTTTTCAGCCACGAGGGTATGAATGCCGGAGTGTCTGTGAACGAGTGTATTAATCTGAA 1713  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
DB 1714 TATTGTACTGAGACTCTGGTCAGTGCCACCAAAATCTTCATTAAG 1758  
RESULT 3  
E26532  
LOCUS E26532 3054 bp DNA linear PAT 18-JUN-2001  
DEFINITION Novel protein belonging to MDC gene family and DNA encoding the  
same.  
ACCESSION E26532  
VERSION E26532.1 GI:13026199  
KEYWORDS JP 1999155574-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3054)  
AUTHORS Kojima, S.  
TITLE Novel protein belonging to MDC gene family and DNA encoding the  
same  
JOURNAL Patent: JP 1999155574-A 3 15-JUN-1999;  
EIGAI CO LTD  
COMMENT OS Homo sapiens (human)  
PN JP 1999155574-A/3  
PD 15-JUN-1999  
PR 01-DEC-1997 JP 1997330020

PI KOJI SONE  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/21, C12P21/02, C12Q1/68//  
(C12N15/09, C12R1:91), (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), PC  
C12N15/00,  
PC (C12N15/00, C12R1:91)  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT CDS 224..2719.  
Location/Qualifiers  
1..3054  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 815 a 754 c 803 g 682 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.12e-30 Length: 3054  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x E26532 (1-3054)  
QY 1 CysSerLeuSerArgnGlyAlaHisCysSerArgPyrCysCysAenAenThrSerCys 20  
Db 1817 TGTTCCTCTCCAAAGCGGCTCAGTCAGCGAGCGGCGCTCTGTAAATAAATCCTCATGT 1876  
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValangluCysAspIleThrGlu 40  
Db 1877 CTTTTCAGCCACGAGGATGATGATGCGGAGATGCTGTGAACGATGTGATATTACTGAA 1936  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
Db 1937 TATTGTACTGGAGACTGTGTGAGTCCAGCCACCAATCTTCATTAAG 1981  
RESULT 4  
AB009672 3054 bp mRNA linear PRI 15-AUG-1998  
LOCUS Homo sapiens mRNA for MDC3, complete cds.  
DEFINITION  
ACCESSION AB009672.1 GI:3419877  
VERSION AB009672.1 GI:3419877  
KEYWORDS MDC3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Sagane, K., Ohya, Y., Hasegawa, Y. and Tanaka, I.  
TITLE Metalloprotease-like, disintegrin-like, cysteine-rich proteins  
MDC2 and MDC3: novel human cellular disintegrins highly expressed  
in the brain  
JOURNAL Biochem. J. 334 (Pt 1), 93-98 (1998)  
MEDLINE 98359734  
PUBMED 963107  
2 (bases 1 to 3054)  
Sagane, K.  
REFERENCE Direct Submision  
TITLE Submitted (11-DEC-1997) Koji Sagane, Eisai Co., Ltd., Tsukuba  
JOURNAL Research Laboratories; Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635,  
Japan (E-mail: k1-sagane@eisai.co.jp, Tel: +81-298-47-5813,  
Fax: +81-298-47-5367)  
Location/Qualifiers  
1..3054  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
224..2722  
/codon\_start=1  
CDS

/product="MDC3"  
/protein\_id="BAA32351.1"  
/db\_xref="GI:3419878"  
/translation="MKPDCSSSRQPLAGCSLAGAGCGPORGAGSVASAPARTTDC  
RLLVLLLPRLAASSRRPARGAASAPAHMETATKKNGLVLDENPTLOQSSNTS  
YSNAQKRETLLEPRLIYINQDESFPYHLDTRARHQKRNKAVHLAQSFOLEAGS  
KFTLLDILNNGLSIDYELHYENGRPOYSKGEGHYGSIIGVDSKVALSTCNGL  
HGFEEDDTFVYMIIEPLVHDEKSTGRPHIIOKTLAAGVYSKQMKNLTMERGDOMPLS  
ELQMLRRKRAVNPSSRGIPEEMKYLEIMIVNDKTYKRNKSHAHNTNFAKSVVNLVD  
SIYKQLNTRVVLVAVETMETEKQDIDITNPVOMLHBFKSYRQRIKOHADAVHLISV  
TFHYKSSLSYFPGVCSRRTRGVNENYGLPMAVQVLSQSLAONLCIQMPSRRKRC  
DCTESWGCTMETETGVSHSRKSKSILETRDVLQGGACLTNRPTKLEPTECGNG  
YVAGEBCDGFVEVCYGLCCCKCSLSNGHSCSDPCNNTSCLFQPRGYECRDVANE  
CDITTEYCTGDSGQCPNHLKQDQVACNONGRCYNEGCKTRDQCQYIMGTKAAGSK  
FCYEKNTGCTEGKNCXKQDRIWIOSKSDVFCGFLCTLTRAPRIGOLOGEIIPTS  
FYHOGRIYDSCGAHVLLDDTDVGYVEDGPGSPMMCCLDKCLQIOLAMSSCPIDS  
KGVCSGHGVCNSAETATCIDFTAGTDCSTRDPVRLNHPKDXGPGSPATNLIGSI  
AGALVAVLGGTGMGFKNVKKRRFPDQGP1"  
BASE COUNT 815 a 754 c 803 g 682 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.12e-30 Length: 3054  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x AB009672 (1-3054)  
QY 1 CysSerLeuSerArgnGlyAlaHisCysSerArgPyrCysCysAenAenThrSerCys 20  
Db 1817 TGTTCCTCTCCAAAGCGGCTCAGTCAGCGAGCGGCGCTCTGTAAATAAATCCTCATGT 1876  
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValangluCysAspIleThrGlu 40  
Db 1877 CTTTTCAGCCACGAGGATGATGATGCGGAGATGCTGTGAACGATGTGATATTACTGAA 1936  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
Db 1937 TATTGTACTGGAGACTGTGTGAGTCCAGCCACCAATCTTCATTAAG 1981  
RESULT 5  
HSAS580 3079 bp mRNA linear PRI 18-JAN-2001  
LOCUS Homo sapiens mRNA for adam23 protein.  
DEFINITION  
ACCESSION AU005580  
VERSION AU005580.1 GI:12053562  
KEYWORDS adam23 gene; Adam23 protein; disintegrin; metalloprotease.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Cal, S. and Lopez-Otin, C.  
TITLE Identification and characterization of ADAM 23, a novel  
metalloprotease/disintegrin protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3078)  
AUTHORS Lopez-Otin, C.  
TITLE Direct Submision  
JOURNAL Submitted (15-APR-1998) Lopez-Otin C., Biocumunica y Biologia  
Molecular, Universidad de Oviedo, C/ Julian Claveria 6. Oviedo.  
Aaturias,., 33006, SPAIN  
REMARK revised by author 23-APR-1998, revised by author 26-OCT-1998  
Location/Qualifiers  
1..3079  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
1..3079  
gene

```
5'UTR /gene="adam23"
CDS <1..45
/gene="adam23"
53..2551
/gene="adam23"
//function="disintegrin"
/codon_start=1
/product="Adam23 protein"
/protein_id="CAC20565.1"
/db_xref="GI:12053563"
/translation="MKPPGSSSRROPPLIAGCSLIGASGCGPGRGAGVAPAPRPNC
RLVLLPLPLAASRRAPRAAAPSAPHNNEETKRLGVADADNTLQONSSNIS
YSNAQKETLPSRLITYINQDSSEPHVLDTKARHQKHNAVLQASFOIEAFGS
KFIIDLINNGLLSSDVEIHYENKGPQISKGEHCYHGSIRGVKDSVALSTNGI
HMFEDDITFVMIIEPLVDEKSTGRPHIIQKTLAQGYKQMKMLTMRGQVPLS
ELMKLRRAVNSRGIPEEMKYLEMLVNDHTYKGRSSHAHTNPAKVAVLVD
STYKOLNRVVLVAVTETKDOIITTPVOMLHPSKYORIKQADAVHLISRV
TTHYKRSLSITFGVCSRTGCVGNBYGLPAAVAVLSQSLQNLGIQWEPSSRKPIC
DCTESWGCGIMEETVSHSRKFSKSIILEYDFLQRGGACLFNRPFLFEPTGNG
VYAGEBCDCGFHVECYGLCKKCSLSNGHSCSDGPCNNNTSCIFQPGYECRDVNE
CDITEYCTGDSGQCPNLHKQDPAONQGRCYNGECKTNDKOCYIWTGTAAGSDK
FCYKLNTEGTEKNGCKGDMIOCSKHDFCGFLCTLNLRAPRIQLOGEIIPIS
FYHOGRVIDCSGAHVLDLDDPDVGVEDGTGCGSMCLDRKCIQIOLNMSCTIDS
KGRVSGHVCSENEATCTICDFTWAGTDCSIRDPVRLNLPKDEGRGFSATNLIIGSI
AGAILVAALVLDGTGMGFKNVKKRRFPDQGPPI"
BASE COUNT 865 a 714 c 769 g 731 t
ORIGIN
3'UTR
/gene="adam23"
/gene="adam23"

Alignment Scores:
Pred. No.: 1,13e-30 Length: 3079
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-634-252A-4_COPY_532_586 (1-55) x HSA5580 (1-3079)
QY 1 CysSerLeuSerArgnGlyAlaHisCysSerArgnGlyProCysCysAsnAsnThrSerCys 20
Db 1646 TTTTCCCTCTCCACAGGGGCTCAGTCAGCAGCGGGCCCTGCTGTACCAATACCTCATGT 1705
QY 21 LeupheginProArgGlyTyrGluCysArgAspAlaValaAngluCysAspIleThrGlu 40
Db 1706 CTTTTCAGCCACAGAGGATGATGATGCCGATGCTGTGAAGAGATGATTTACTGAA 1765
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db 1766 TATTGTACTGAGAGACTGTGTGAGTCCACCAAAATCTTCATTAAG 1810

RESULT 6
LOCUS AX299710 4043 bp DNA linear PAT 26-NOV-2001
DEFINITION Sequence 2 from Patent WO0174857.
ACCESSION AX299710
VERSION AX299710.1 GI:17129252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lopez-Otin,C., Freilij,J.M., Bianchi,A.B., Muguel,S.C., Garcia,J.M.
and Trall,P.
TITLE Methods and compositions for modulating integrin-mediated cell-cell
interactions
JOURNAL Patent: WO 0174857-A 2 11-OCT-2001;
Bristol-Myers Squibb Co. (US)
FEATURES
source 1..4043
/organism="Homo sapiens"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1110 a 946 c 1002 g 985 t
ORIGIN
Alignment Scores:
Pred. No.: 1.47e-30 Length: 4043
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-634-252A-4_COPY_532_586 (1-55) x AX299710 (1-4043)
QY 1 CysSerLeuSerArgnGlyAlaHisCysSerArgnGlyProCysCysAsnAsnThrSerCys 20
Db 2610 TTTTCCCTCTCCACAGGGGCTCAGTCAGCAGCGGGCCCTGCTGTACCAATACCTCATGT 2669
QY 21 LeupheginProArgGlyTyrGluCysArgAspAlaValaAngluCysAspIleThrGlu 40
Db 2670 CTTTTCAGCCACAGAGGATGATGATGCCGATGCTGTGAAGAGATGATTTACTGAA 2729
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db 2730 TATTGTACTGAGAGACTGTGTGAGTCCACCAAAATCTTCATTAAG 2774

RESULT 7
LOCUS AB009673 2891 bp mRNA linear ROD 17-AUG-1999
DEFINITION Mus musculus mRNA for ADAM23, complete cds.
ACCESSION AB009673
VERSION AB009673.1 GI:5736618
KEYWORDS ADAM23.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Sagane,K., Yamazaki,K., Mizui,Y. and Tanaka,I.
TITLE Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23
JOURNAL Gene 236 (1), 79-86 (1999)
MEDLINE 99365303
PUBMED 10433968
REFERENCE 2 (bases 1 to 2891)
AUTHORS Sagane,K.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1997) Koji Sagane, Eisai Co., Ltd., Tsukuba
Research Laboratories, Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635,
Japan (E-mail:k-sagane@hcc.eisai.co.jp, Tel:+81-298-47-5813,
Fax:+81-298-47-5367)
COMMENT Sequence updated (07-Aug-1999).
FEATURES
source Location/Qualifiers
1..2891
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
60..2569
/codon_start=1
/product="ADAM23"
/protein_id="BAA63381.1"
/db_xref="GI:5736619"
/translation="MKPPGSSSRROPPLIAGCSLIGASGCGPGRGAGVAPAPRPNC
RLVLLPLPLAASRRAPRAAAPSAPHNNEETKRLGVADADNTLQONSSSRNTSYSS
AVQKEITDPSRLVITYINQDSSEPHVLDTKARHQKHNAVLQASFOIEAFGSKFI
LDLTLNNGLLSSDVEIHYEDGKQMYSGGHCYHGSIRGVKDSVALSTNGIHM
FEDDITFVMIIEPLVDEKSTGRPHIIQKTLAQGYKQMKMLTMRGQVPLS
ELMKLRRAVNSRGIPEEMKYLEMLVNDHTYKGRSSHAHTNPAKVAVLVD
STYKOLNRVVLVAVTETKDOIITTPVOMLHPSKYORIKQADAVHLISRVTH
YKRSLSITFGVCSRTGCVGNBYGLPMAVAVLSQSLQNLGIQWEPSSRKPICECI
ESWGCGIMEETVSHSRKFSKSIILEYDFLQRGGACLFNRPFLFEPTGNGYVE
AGBECDCGFHVECYGCKKCSLSNGHSCSDGPCNNNTSCIFQSRGYECRDVANS CDI
```

polvA\_s1site  
1450  
1295  
TEYCTGAGGACCPYLNKDDGYSANONORCAVNGECKTRDNOCCYIIGWKAAGSRFCY  
EKTNEGTEKSGCGCKDGDRMI PCSKHDIYPCGGLICTNHTRAPRITGOLOGEIIPSTFTH  
OGRVIDSGAHVLLDDTDVGEDVDYPCGSGMMCLNKKCQIQIQLNNSSCPLDSRGK  
VCSGAVSSENEATCICDPTWAGTDCSIDPVPANPNPFIKDEBPGKPSATNLIIIGSIAAG  
ILVAIVLIVGCTGWFKNVKKSRFDPDQGGP1

BASE COUNT	/note="10 a nucleotides"		
ORIGIN	763 a	747 c	763 g 618 c

Alignment Scores:	
Pred. No.:	4, 1e-29
Score:	329.00
Percent Similarity:	96.36
Best Local Similarity:	96.36
Query Match:	96.08
R.	10
Gene.	0

US-09-634-252A-4\_COPY\_532\_586 (1-55) X AB009673 (1-2891)

QY 1 CysSerLeuSerAngIValAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20  
Db 1664 TGTTGCTCTCCATGGGGCCACTGCAGTACGGCCCCCTGCTETHAACACACCTCATGT 1733

Oy 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
 Db 1724 CTTTTTCACTCAGGAGTATGATGTCGGATGCGTAAACACTGTGTATATACACCGAG 1783

QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55  
DB 1784 TACTGCACTGGAGACTCTGGCCAGTCCCAACCGAAGCTCCATAA 1328

LOCUS	E10490	1464 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	cDNA encoding a consensus region of MDC protein.				
FEATURES	<p>1..1464</p>				

VERSION	E10490.1	GI:22027323
KEYWORDS	JP 1995330799-A/1.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 1464)  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
Nakamura, Y. and Emi, M.  
MDC PROTEIN AND DNA CODING THE SAME  
Patent: JP 1995330799-A 1 19-DEC-1995;

**COMMENT**

PD 19-DEC-1995 JP 1594084470  
 PD 22-APR-1994 JP 153602, 22-SEP-1993, JP 93P 257455, PR  
 PR 14-MAY-1993 JP 93P 49804, 12-APR-1994 JP 94P 73328 PI  
 23-FEB-1994 JP 94P  
 NAKAMURA YOSUKU, ENMI MITSURU  
 C07K16/32, C07K14/82, C12N1/21, C12N15/02, C12N15/09, C12P21/02, PC  
 C12P21/08  
 PC C12P21/08, GO1N33/53, GO1N33/574, (C12N1/21, C12R:119), (C12N15/09,  
 C12P21/08)

PC	(C12P21/02,C12R1:19), (C12P21/08,C12R1:91),
CC	strandness: Double;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
EH	Key Location/Qualifiers

```
FT source
FT 1. .1464
FT /organism='Homo sapiens.'
FT /tissue_type='Cerebellum
CDS <1. .>1464
FT /product='MDC protein'.
FT
```

**FEATURES**  
**source**

```

/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
BASE COUNT      294 a      445 c      473 g      253 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	1,87e-16
Score:	213.00
Percent Similarity:	76.364
Best Local Similarity:	65.458
Query Match:	64.168
DB:	6
	Gaps: 1
	Length: 146
	Matches: 36
	Conservative: 6
	Mismatches: 11
	Indels: 2

US-09-634-252A-4\_COPY\_532\_586 (1-55) X E10490 (1-1464)

Oy 1 CysSerLeuSerAngLYAlaHisCysSerAsnGlyProCysCysAsnAsnThrSerCys 20  
 1111 TGCACCTGACTACGACCCCAATGTGACAGCGAGGCTTGTCTTTCGCGGC-----TGC 1164

QY 21 LeudheginProxrGGLTYrGrucCysArAsplAValAlasnrcuCysAspiIethrcju 40  
:::.....  
Db 1165 AAGTACGAACCACCGGGGTGTCTCTCCGACGAGGCCGTAAACGCATGCCATTCGGGAG 1224

QY      41 TyrCysThrGlyAspSerGlyIncysProProAsnLeuHisLys 55  
         |||||  
Db    1225 ACCTGCACCGGGACTCTAGCCAGTGCCCGCTAACTGCACAAG 1265

RESULT 9			
125845			
LOCUS	125845	1464 bp	DNA
DEFINITION	Sequence 5 from patent US 555526.		linear
			PAT 07-OCT-1996

VERSION	125845.1	GI.1605715
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
	Unclassified.	

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1 (bases 1 to 1464)	Nakamura, Y. and Emi, M.	MDC proteins and DNAs encoding the same	Patent: US 552526-A 5 03-SEP-1996;	Location/Qualifiers	1..1464

	/organism="unknown"			
BASE COUNT	294 a	445 c	473 g	252 t
ORIGIN				

**Alignment Scores:**

Pred. No.:	1,87e-16	Length:	1466
Score:	213.00	Matches:	36
Percent Similarity:	76.36%	Conservative:	6
Best Local Similarity:	65.45%	Mismatches:	11
Query Match:	64.16%	Indels:	2
DB:	6	Gaps:	1

1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyPro

Db 1111 TGCACCTGACTCAGCACCAGTCATGTGCACGACCGGCTCTGCTGTGCGCCGC-----TGC 1164

Db 1165 AAGTACGAACACCGGGGTGTGTCCTGCAGAGAGCCGTGAACGATGCCACATCGCGAG 1234

0y 41 TycCstTrtG1yAaPSeG1yGlnCysaProkAnseuH1slyv 55  
Db 1225 ACtGcACcGgGgAcTtAGcCAgTgcCGcCTAAcTGCACAAg 12

I43482	1464 bp	DNA	linear	PAT 07-OCT-1997
LOCUS	I43482			

```

DEFINITION Sequence 5 from patent US 5631351.
ACCESSION 143482
VERSION 143482.1 GI:2468726
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1464)
TITLE Nakamura, Y. and Emi, M.
JOURNAL Antibodies to MDC proteins
PATENT Patent: US 5631351-A 5 20-MAY-1997;
FEATURES
    source
        1..1464
        /organism="unknown"
BASE COUNT 294 a 445 c 473 g 252 t
ORIGIN

Alignment Scores:
Pred. No.: 1.87e-16 Length: 1464
Score: 213.00 Matches: 36
Percent Similarity: 76.36% Conservative: 6
Best Local Similarity: 65.45% Mismatches: 11
Query Match: 64.16% Indels: 2
DB: 6 Gaps: 1

US-09-634-252a-4_COPY_532_586 (1-55) x 143482 (1-1464)
QY 1 CYSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20
   |||:::|||||:::|||||
Db 1111 TGCACCTTACTACTACGACGCGCATGTGACGAGCGGCTCTGCTGTGCGCCG-----TGC 1164

QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
   :::::|||||:::|||||
Db 1165 AAGTACGAACCAACGGGGTGTGTCTGCGCGAGAGCGCGCTGTGACGAGTGGCATCGCGGAG 1224

QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
   |||:::|||||:::|||||
Db 1225 ACCTGCACCGGGGACTCTAGCGAGTGTCCGCGCTTAACCTGCACAG 1269
  
```

Search completed: October 21, 2003, 12:02:01  
 Job time : 1821.35 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:11:41 ; Search time 142.862 Seconds  
(without alignments)  
1039.251 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CUSLNGAHCSDGPCNNNTSC.....CDITEYCTGDSGQCPNLMK 55

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Delop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1h  
-Q=/cgrn\_1/USFO.spool/US09634252/runat\_21102003\_090322\_3362/epq\_query.fasta\_1.462  
-DB=Geneseq\_19Jun03 -Qfmt=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=10  
-MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09634252 @C31 1.401 @runat\_21102003\_090322\_3362 -NCPU=6 -ICPU=3  
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: Geneseq\_19Jun03 : \*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1960.DAT : \*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1961.DAT : \*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1962.DAT : \*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1963.DAT : \*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1964.DAT : \*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1965.DAT : \*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1967.DAT : \*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1968.DAT : \*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1969.DAT : \*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT : \*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT : \*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT : \*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT : \*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT : \*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT : \*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT : \*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT : \*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT : \*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT : \*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT : \*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT : \*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT : \*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2004.DAT : \*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2005.DAT : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	332	100.0	1668	22	AA021439	ADAM-23dis-Fc fusi
2	332	100.0	2268	21	AA245836	DNA encoding a dis
3	332	100.0	2268	22	AA511991	Human CDNA encodin
4	332	100.0	2268	24	AB078447	Nucleotide sequenc
5	332	100.0	2499	20	AA210207	CDNA encoding the
6	332	100.0	3054	20	AA210207	Human MDC3 CDNA.
7	332	100.0	3054	25	ABX76201	Lung cancer-associ
8	332	100.0	3054	25	ABX76202	Lung cancer-associ
9	332	100.0	4043	24	AA172025	ADAM 23 CDNA. Hom
10	285	85.8	2088	21	AA245837	Degenerate DNA enc
11	285	85.8	2088	22	AA511992	Human degenerate D
12	285	85.8	2088	24	AB078448	Degenerate nucleot
13	213	64.2	1464	16	AA076119	Human fetal cerebe
14	213	64.2	2913	16	AA076121	Human fetal cerebe
15	213	64.2	2913	16	AA076120	Human fetal cerebe
16	213	64.2	3183	16	AA076122	Human fetal brain
17	185	55.7	1674	22	AA021438	ADAM-22dis-Fc fusi
18	185	55.7	2604	20	AA210208	CDNA encoding the
19	185	55.7	2697	20	AA210208	Human MDC2-beta cD
20	185	55.7	2805	20	AA210208	Human MDC2-beta cD
21	173.5	52.3	2653	17	AA174616	CRII-7 nerve prote
22	168.5	50.8	276	25	AB268563	Nucleotide sequenc
23	168.5	50.8	1638	22	AA021434	ADAM-15dis-Fc fusi
24	168.5	50.8	2720	24	AB235307	Human gene express
25	168.5	50.8	2740	24	AB235307	Human gene express
26	168.5	50.8	2740	24	AB235307	Breast cancer-asso
27	168.5	50.8	2740	24	AB235307	Gene #3360 used to
28	168.5	50.8	2968	22	AA514756	Human protease cDN
29	167.5	50.8	2968	24	AB414523	Human metallopept
30	167.5	50.8	1500	21	AA459303	DNA encoding a sna
31	167.5	50.3	2346	21	AA459308	DNA encoding a sna
32	167.5	50.3	2373	21	AA459308	CDNA encoding a hu
33	167.5	50.3	2373	21	AA459307	DNA encoding a sna
34	167.5	50.3	2695	22	AA51345	Human cDNA sequenc
35	167.5	50.3	2695	25	ACA03704	CDNA encoding huma
36	167.5	50.3	2695	25	ACA03704	Human cDNA encodin
37	167.5	50.3	2695	25	ABX89242	DNA encoding novel
38	162	48.8	2058	23	AB141903	Drosophila melanog
39	160	48.2	1507	24	ABK30352	Human G-protein-co
40	160	48.2	2642	24	ABT08500	Human novel protei
41	160	48.2	2649	24	AA597181	Human metallopept
42	160	48.2	2705	24	ABT08499	Human novel protei
43	160	48.2	2762	24	ABT08498	Human novel protei
44	160	48.2	3015	23	AA517731	DNA encoding novel
45	160	48.2	3431	22	AA05513	First splice varia

## ALIGNMENTS

RESULT 1  
AAD21439  
ID AAD21439 standard; DNA, 1668 BP.  
XX  
AC AAD21439;  
XX  
DT 28-JAN-2002 (first entry)  
XX  
DE ADAM-23dis-Fc fusion construct DNA.  
XX  
Human; ADAM disintegrin domain; integrin; endothelial cell migration;  
XX angiogenesis; ocular disorder; inflammatory disease; bone resorption;  
XX osteoporosis; retinosis; thrombosis; tissue repair; wound healing;  
XX retinopathy; retinoblastoma; neovascular glaucoma; macular degeneration;  
XX retrolental fibroplasia; inflammatory bowel disease; rubeosis; uveitis;  
XX arthritis; rheumatism; myocardial infarction; coronary artery disease;  
XX tumour; stroke; atherosclerosis; arteriosclerosis; injury; psoriasis;

CC	artery disease, or for occlusion, reocclusion, stenosis and/or restenosis
CC	of blood vessels or stroke benign tumours and preneoplastic conditions,
CC	myocardial angiogenesis, haemophilic joints, scleroderma, vascular
CC	adhesions, asthma and allergy, eczema and dermatitis, graft versus host
CC	disease, sepsis, adult respiratory distress syndrome, telangiectasia and
CC	wound granulation. The method are used in combination with angioplasty
CC	procedures, such as balloon angioplasty, laser angioplasty, coronary
CC	atherectomy or similar techniques, carotid endarterectomy, anastomosis of
CC	vascular grafts, surgery having a high risk of thrombus formation (i.e.
CC	coronary bypass surgery, insertion of a prosthetic valve or vessel and
CC	the like), atherectomy, stent placement, placement of a chronic
CC	cardiovascular device such as an in-dwelling catheter or prosthetic valve
CC	or vessel, organ transplantation or bypass surgery. The present sequence
CC	is a DNA encoding ADAM disintegrin domain polypeptide fusion construct.
CC	The fusion construct comprises of immunoglobulin K (Igk) leader, human
CC	ADAM disintegrin and Fc region.
XX	
SO	Sequence 1668 BP; 423 A; 444 C; 455 G; 346 T; 0 other;
Alignment Scores:	
Pred. No.:	2,71e-30
Score:	332.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	22
US-09-634-252A-4_COPY_532_586 (1-55) x AAD21439 (1-1668)	
Qy	1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAenThrSerCys 20
Db	172 TGTTCCTCTCCACGCGGCTCAGTCAGCAGCGGCGCTCTTAACATACCTCATGT 231
Qy	21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValaGlnGluCysAspIleThrGlu 40
Db	232 CTTTTCACGCCAGAGGGATGAATGCCGGATGCTGTGAACGAGTGGATTACTGAA 291
Qy	41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIshs 55
Db	292 TATTGTACTGGAGACTCTGTGTCAGTCCACCAATCTTATTAAAG 336
RESULT 2	
AAZ45826	
ID	AAZ45826 standard; DNA; 2268 BP.
XX	
AC	AAZ45826;
XX	
DT	25-APR-2000 (first entry)
DE	DNA encoding a disintegrin homologue designated zdint1.
XX	
KW	Human; disintegrin homologue; zdint1, cardiac myocyte; adipocyte;
KW	gene therapy; cell-cell interaction; chondro sacroma; atherosclerosis;
KW	Alzheimer's disease; restenosis; ischemic reperfusion; obesity;
KW	intimal hyperplasia; tumour; platelet aggregation; apoptosis;
KW	neutogenesis; myogenesis; arthritis; myogenesis; neurogenesis;
KW	connective tissue disorder; chondrogenesis; tumour proliferation;
KW	inflammation; ss.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	CDS
FT	Location/Qualifiers
FT	3..2093
FT	/*tag= a
FT	/product= "zdint1"
XX	
PN	WO200002912-A2.
XX	
PD	20-JAN-2000.
XX	
PF	09-JUL-1999; 99MO-US15638.
XX	
RR	10-JUL-1998; 98US-0113883.



RESULT 4  
 ID ABO78447 standard; cDNA; 2268 BP.  
 XX ABO78447;  
 AC  
 DT 05-NOV-2002 (first entry)  
 XX  
 XX Nucleotide sequence of human zdn1.  
 XX  
 XX Human; zdn1; disintegrin protease; platelet accumulation;  
 KW chromosome 2q33; platelet aggregation; proteolysis; apoptosis;  
 KW neurogenesis; myogenesis; connective tissue disorder; arthritis;  
 KW chondrogenesis; cell adhesion; cell fusion; neuron; myocyte; heart;  
 KW brain; spinal cord; skeletal muscle tissue; Alzheimer's disease;  
 KW tumour formation; multiple sclerosis; congestive heart failure;  
 KW ischaemic reperfusion; intimal hyperplasia; restenosis; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 3..2093  
 FT CDS /\*tag= a  
 FT /\*product= "zdn1"  
 XX  
 XX US2002072102-A1.  
 PN  
 XX  
 PD 13-JUN-2002.  
 XX  
 PF 16-MAR-2001; 2001US-0809790.  
 XX  
 PR 10-JUL-1998; 98US-092371P.  
 PR 09-JUL-1999; 99US-0351414.  
 XX  
 PA (SHEP/) SHEPPARD P O.  
 PA (BAIN/) BAINDUR N.  
 PA (DEIS/) DEISHER T A.  
 PA (BISH/) BISHOP P D.  
 XX  
 PI Sheppard PO, Bainsur N, Deisher TA, Bishop PD;  
 XX  
 DR WPI; 2002-598452/64.  
 DR P-PSDB; ABB78130.  
 XX  
 PT Modulating cell-cell interactions and diagnosis, treatment of  
 PT Alzheimer's disease, tumour formation, multiple sclerosis and congestive  
 PT heart failure -  
 XX  
 XX Claim 14; Page 26-29; 53pp; English.  
 XX  
 XX The present sequence encodes a human polypeptide designated zdn1.  
 CC zdn1 is a member of the disintegrin protease family; zdn1 inhibits  
 CC platelet accumulation. The zdn1 gene is present on chromosome 2q33.  
 CC zdn1 polypeptides and polynucleotides are useful in treatment of  
 CC disorders associated with infarct in brain or heart tissue and/or  
 CC platelet aggregation, to modulate proteolysis, apoptosis, neurogenesis,  
 CC myogenesis, connective tissue disorders, arthritis, chondrogenesis,  
 CC cell adhesion, cell fusion, and signalling or to treat or prevent  
 CC development of pathological conditions in such diverse tissue as heart,  
 CC brain, spinal cord and skeletal muscle. The molecules modulate  
 CC inhibition and proliferation of neurons and myocytes in heart, brain,  
 CC spinal cord and skeletal muscle tissue. Disorders which may be amenable  
 CC to diagnosis, treatment or prevention with zdn1 polypeptides include,  
 CC Alzheimer's disease, tumour formation, multiple sclerosis, congestive  
 CC heart failure, ischaemic reperfusion or infarct and degenerative  
 CC diseases. The zdn1 molecules particularly useful in the treatment of  
 CC intimal hyperplasia or restenosis due to acute vascular injury.  
 XX  
 XX Sequence 2268 BP; 654 A; 457 C; 558 G; 598 T; 1 other;  
 XX  
 Alignment Scores: 3.95e-30 Length: 2268  
 Pred. No.:

	Score:	332.00	Matches:	55
	Percent Similarity:	100.00%	Conservative:	0
	Best Local Similarity:	100.00%	Mismatches:	0
	Query Match:	100.00%	Indels:	0
	DB:	24	Gaps:	0
US-09-634-252A-4_COPY_532_586 (1-55) x ABQ78447 (1-2268)				
OY	1	CysSerLeuSerArgAngLYAlaHisCysSerAspGLyProCYsCYsAsnAsnThrSerCYs	-20	
Dd	1227	TGTTCCCTCTTCACAGGGGGCTCACTGCAGCACGGGCCCTCTGTAAACAATACCCTCATGT	1286	
OY	21	LeuHeGInPcOAvggIYTyrGlUcYAArgspAlaValaLaangLUcYsASpILlethrGLU	40	
Dd	1287	CTTTTTCAGCCACGAGGGTAAGAATGCCGGGATCTGTGAACGAGTGATATTACTGAA	1346	
OY	41	TyrCYSThrGIYAASPserGLyglUCysProPCoAsnLeuHISlys	55	
Dd	1347	TATTGTACTGAGACTCTGTGTCAGTCCACCAAATCTTCAATAAG	1391	
RESULT 5				
AAZ10207	ID	AAZ10207 standard; DNA; 2499 BP.		
XX	AC	AAZ10207;		
XX	DT	29-OCT-1999 (first entry)		
XX	DE	cDNA encoding the human SYPH3-17 protein.		
KW	SYPH3-17; metalloproteinase-disintegrin; human chromosome 2;			
KM	human chromosome 7; malignant hypertension susceptibility;			
KW	Zellweger syndrome; neonatal adrenoleukodystrophy;			
KW	infantile Reifsum disease; progressive familial intrahepatic cholestasis;			
KW	mucopolysaccharidosis VII; split hand/foot malformation;			
KW	arhythmogenic right ventricular dysplasia-4; Coppock-like cataract;			
KW	insulin dependent diabetes mellitus-12; lamellar type ichthyosis;			
KW	transient neonatal myasthenia gravis; congenital aculeiform cataract;			
KW	juvenile amyotrophic lateral sclerosis;			
KW	familial paroxysmal choreoathetosis;			
KW	Finnish lethal neonatal metabolic syndrome; ss.			
XX	Homo sapiens.			
XX	OS			
XX	Key	Location/Qualifiers		
FX	FT	1..2499		
XX	CDS	/*tag= a		
XX	WT			
XX	FN	WO9941388-A2.		
XX	PD	19-AUG-1999.		
XX	PF	11-FEB-1999; 99WO-US03016.		
XX	PR	11-FEB-1998; 98US-0074310.		
PA	(IMMV ) IMMUNEX CORP.			
PI	Cerrecti DP;			
DR	WPI; 1999-527371/44.			
PT	P-PsDB; AAY30207.			
XX	DNA encoding the SYPH3-13 and SYPH3-17 proteins for detecting			
XX	disease corresponding to chromosome 7, e.g. Zellweger syndrome			
XX	Claim 1; Page 7-8; 82pp; English.			
CC	The present sequence encodes a protein designated SYPH3-17, which			
CC	is a member of the metalloproteinase-disintegrin family. The			
CC	specification also describes SYPH3-13 proteins. Both sequences can			
CC	be used to identify human chromosome 2 or 7, and to map genes on these			
CC	two chromosomes, and also to identify genes associated with certain			

CC diseases, syndromes, or other human conditions associated with human  
CC chromosome 2 or 7. The disease that correspond to chromosome 7 include  
CC malignant hyperthermia susceptibility, Zellweger syndrome, neonatal  
CC adrenoleukodystrophy, infantile Refsum disease, progressive familial  
CC intrapapillary colobomata, mucopolysaccharidosis VII, and split  
CC hand/foot malformation. Diseases associated with chromosome 2 include  
CC arylsulphatase B deficiency, congenital myasthenia gravis, juvenile amyotrophic  
CC lateral sclerosis, congenital aculeiform cataract, Cypock-like cataract,  
CC lamellar type ichthyosis, familial paroxysmal choreoathetosis, and  
CC Finnish lethal neonatal metabolic syndrome.

XX SQ Sequence 2499 BP; 677 A; 593 C; 645 G; 584 T; 0 other;

Alignment Scores:

Pred. No.: 4,466-30 Length: 2499  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x AA210207 (1-2499)

OY 1 CysSerLeuSerAenGlyAlaHisCysSerAenGlyProCysCysAenAenThSerCys 20  
Db 1594 TGTTCCTCTCCAAAGGGGCTCAGTCAGCGAGCGGCTCTGTAAACAATACCTCATGT 1653

OY 21 LeuPheGlnProAenGlyTyrGluCysArgAspAlaValAenGluCysAsp11ethrGlu 40  
Db 1654 CTTTTCAGCCACGAGGATGAATGCCGGATGCTGGAAGAGTGTGATATTACTGAA 1713

OY 41 TyrCysThrGlyAspSerGlyGlnCysProProAenLeuHisLys 55  
Db 1714 TATTGTACTGAGACTCTGTCTAGTCCACCAATCTTCATTAA 1758

RESULT 6

AAx78438 standard; cDNA to mRNA; 3054 BP.

XX AC AAX78438;  
XX DT 26-AUG-1999 (first entry)  
XX DE Human MDC3 cDNA.  
XX KW Metalloproteinase-like-disintegrin-like cysteine rich protein; human;  
XX KM MDC2-alpha; MDC2-beta; MDC3; medical; treatment; diagnosis; ds.  
XX OS Homo sapiens.

Location/Qualifiers

Key 224..2722  
FT /\*tag= a  
FT /product= "MDC3"

JP1155574-A.

15-JUN-1999.

01-DEC-1997; 97JP-0330020.

01-DEC-1997; 97JP-0330020.

(EISA ) EISA CO LTD.

WPI; 1999-398071/34.

P-PSDB; AAY25120.

New protein belonging to MDC gene family - useful in medical

treatment and diagnosis

XX XX This invention describes novel human MDC2-alpha, MDC2-beta and MDC3  
CC proteins and their encoding nucleic acids. The products of the invention  
CC are useful in medical treatment and diagnosis.  
XX SQ Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other;

Alignment Scores:

Pred. No.: 5,76-30 Length: 3054  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x AAX78438 (1-3054)

OY 1 CysSerLeuSerAenGlyAlaHisCysSerAenGlyProCysCysAenAenThSerCys 20  
Db 1817 TGTTCCTCTCCAAAGGGGCTCAGTCAGCGAGCGGCTCTGTAAACAATACCTCATGT 1876

OY 21 LeuPheGlnProAenGlyTyrGluCysArgAspAlaValAenGluCysAsp11ethrGlu 40  
Db 1877 CTTTTCAGCCACGAGGATGAATGCCGGATGCTGGAAGAGTGTGATATTACTGAA 1936

OY 41 TyrCysThrGlyAspSerGlyGlnCysProProAenLeuHisLys 55  
Db 1937 TATTGTACTGAGACTCTGTCTAGTCCACCAATCTTCATTAA 1981

RESULT 7

ABx76201 standard; DNA; 3054 BP.

XX AC ABX76201;  
XX DT 02-APR-2003 (first entry)  
XX DE Lung cancer-associated polynucleotide #72.  
XX KW Lung cancer-associated polynucleotide; gene; ds; cytosolic; emphysema;  
XX KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

WO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US12476.

18-APR-2001; 2001US-284770P.

10-MAY-2001; 2001US-290492P.

09-NOV-2001; 2001US-339245P.

13-NOV-2001; 2001US-350665P.

29-NOV-2001; 2001US-334370P.

12-APR-2002; 2002US-372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Aziz N, Murray R;

WPI; 2003-093161/08.

P-PSDB; ABUS6479.

Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
PT expression in lung cancer -  
XX  
XX  
PS Claim 22; Page 243; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung  
CC cancer-associated polynucleotides and polypeptides are used for  
CC identifying a compound that modulates a lung cancer-associated  
CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
CC cell to treat lung cancer in a patient and for treating a mammal having  
CC lung cancer by administering a modulatory compound identified. The  
CC methods are useful for treating lung cancer, such as small cell lung  
CC cancer, non-small cell lung cancer or other benign or precancerous  
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
CC and polypeptides are useful for diagnostic purposes and as targets for  
CC screening for therapeutic compounds that modulate lung cancer, such as  
CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated  
CC polynucleotides of the invention.

XX SQ Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other;

Alignment Scores:  
Pred. No.: 5.7e-30 Length: 3054  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x ABX76201 (1-3054)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysGlyAsnAsnThrSerCys 20  
DB 1817 TGTTCCTCTCCACAGGGGCTCAGTCAGCGAGCGGCCCTGCTGTAACATTCCTATCT 1876

QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
DB 1877 CTTTTCACGCCAGAGGATGATGATCCGGGATGCTGTACAGAGTGATATTACTGAA 1936

QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIys 55  
DB 1937 TATTGACTGGAGACTCTGTCAGTCCACCAATCTTCATTAAG 1981

RESULT 8

ABX76292  
ID ABX76292 standard; DNA; 3054 BP.

XX AC ABX76292;

XX DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #156.

XX Lung cancer-associated polynucleotide; gene; ds; cytosolic; emphysema;  
XX antiinflammatory; antiasbestatic; non-small cell lung cancer; atelectasis;  
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US12476.

XX PR 18-APR-2001; 2001US-284770P.

XX PR 10-MAY-2001; 2001US-280492P.

XX PR 09-NOV-2001; 2001US-339245P.

XX PR 13-NOV-2001; 2001US-350666P.

PR 29-NOV-2001; 2001US-334370P.

PR 12-APR-2002; 2002US-372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX DR WPI; 2003-093161/08.

XX DR P-PSDB; ABUS6563.

PT Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer

PS Claim 22; Page 307-308; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung  
CC cancer-associated polynucleotides and polypeptides are used for  
CC identifying a compound that modulates a lung cancer-associated  
CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
CC cell to treat lung cancer in a patient and for treating a mammal having  
CC lung cancer by administering a modulatory compound identified. The  
CC methods are useful for treating lung cancer, such as small cell lung  
CC cancer, non-small cell lung cancer or other benign or precancerous  
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
CC and polypeptides are useful for diagnostic purposes and as targets for  
CC screening for therapeutic compounds that modulate lung cancer, such as  
CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated  
CC polynucleotides of the invention.

XX SQ Sequence 3054 BP; 815 A; 754 C; 803 G; 682 T; 0 other;

Alignment Scores:  
Pred. No.: 5.7e-30 Length: 3054  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x ABX76292 (1-3054)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysGlyAsnAsnThrSerCys 20  
DB 1817 TGTTCCTCTCCACAGGGGCTCAGTCAGCGAGCGGCCCTGCTGTAACATTCCTATCT 1876

QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
DB 1877 CTTTTCACGCCAGAGGATGATGATCCGGGATGCTGTACAGAGTGATATTACTGAA 1936

QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIys 55  
DB 1937 TATTGACTGGAGACTCTGTCAGTCCACCAATCTTCATTAAG 1981

RESULT 9

AAI72025  
ID AAI72025 standard; cDNA; 4043 BP.

XX AC AAI72025;

XX DT 04-MAR-2002 (first entry)

XX DE ADAM 23 cDNA.

XX KW Disintegrin; ADAM 23; metalloproteinase; domain; adhesion; proteolysis;  
modulator; alpha-v-beta3 integrin; tumour progression; neural tissue;

KM		angiogenesis; ss.
XX	Homo sapiens.	
OS		
XH	Key	Location/Qualifiers
FH	CDS	1017..3515
FT		/tag= a
FT		/product= "ADAM 23"
PX		
PN	W0200174857-A2.	
XX		
XX	11-OCT-2001.	
PD		
XX		
PF	02-APR-2001; 2001WO-US10729.	
XX		
PR	03-APR-2000; 2000US-194164P.	
XX		
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.	
PI	Lopez-Otin C, Freiji JMP, Bianchi AB, Miguel SC, Garcia JML;	
PI	Tail P;	
XX		
DR	WPI; 2002-066298/09.	
DR	P-PSDB; AAB47778.	
PT	Nucleic acid encoding a new ADAM family member, designated ADAM 23 is	
PT	useful to find modulators of its interaction with integrin which can be	
PT	used to prevent angiogenesis or increase neural growth.	
XX		
PS	Claim 2; Page 36-38; 44pp; English.	
CC	This sequence encodes a cellular disintegrin, ADAM 23, which is a	
CC	protein having a disintegrin and metalloprotease domain. This	
CC	protein can perform both adhesion and proteolysis functions.	
CC	Modulators of the interaction between ADAM and alpha-v-beta3 integrin	
CC	are used to inhibit tumour progression or induce growth of neural	
CC	tissue. These compounds modulate angiogenesis and induction of matrix	
CC	metalloproteinases facilitating migration of tumour cells and growth of	
CC	neural tissue.	
XX		
SQ	Sequence 4043 BP; 1110 A; 946 C; 1002 G; 985 T; 0 other;	
Alignment Scores:		
Pred. NO.:	8.06e-30	Length: 4043
Score:	332.00	Matches: 55
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	24	Gaps: 0
•		
US-09-634-252A-4_COPY_532_586 (1-55) x AA172025 (1-4043)		
OY	1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys	20
Db	2610 TGTTTCCTCTTCAACGGGGCTCACCTGCAGCGGCCCTCTGTATACATACTCATGT	2669
OY	21 LeuPheGlnProArgIYTYrGlucysArGaspAlavalasngluCysAspDIEThrglu	40
Db	2670 CTTTTTGCCACGACGAGGTATGAATGCCGGATGTGTGAACGAGTGTATTAAGA	2729
OY	41 TYrCYsrhrgIyaSPserGIyGLInCysePrOProasnLeuhISLyS	55
Db	2730 TATTGTACTGAGACTCTGTGTCAGTCCACCACAATCTTCATTAAG	2774
RESULT 10		
ID	AA245827 standard; DNA; 2088 BP.	
XX		
AC	AA245827;	
XX		
DT	25-APR-2000 (first entry)	
XX		
DE	Degenerate DNA encoding a disintegrin homologue designated zdintl.	

XX	Human; disintegrin homologue; zidm1; cardiac myocyte; adipocyte;
KX	gene therapy; cell-cell interaction; chondro sarcoma; atherosclerosis;
KM	Alzheimer's disease; restenosis; ischemic reperfusion; obesity;
KW	intimal hyperplasia; tumour; platelet aggregation; apoptosis;
KV	neurogenesis; myogenesis; arthritis; myogenesis; neurogenesis;
KW	connective tissue disorder; chondrogenesis; tumour proliferation;
KM	inflammation; ss.
XX	
OS	Synthetic.
XX	Homo sapiens.
XX	
PM	WO200002912-A2.
PD	20-JAN-2000.
PF	
XX	09-JUL-1999; 99WO-US15638.
PR	10-JUL-1998; 98US-0113883.
XX	
PA	(ZYMO ) ZYMOGENETICS INC.
PI	Shepard PO, Balindur N, Delsheer TA, Bishop PD;
XX	WP1; 2000-160896/14.
DR	
XX	Polypeptide useful in modulating cell-cell interaction in tissues of
PT	heart, brain, spinal cord and treating chondro sarcoma,
XX	atherosclerosis, restenosis, obesity, intimal hyperplasia and tumors -
XX	Disclosure; Page 129-130; 132pp; English.
XX	
XX	The present sequence represents a degenerate sequence which encodes
CC	a human disintegrin homologue, designated zidm1. The zidm1 polypeptide
CC	is a cardiac myocyte proliferation and differentiation stimulator, as
CC	well as an adipocyte proliferation and differentiation inhibitor.
CC	Polynucleotides encoding zidm1 are used in gene therapy. The zidm1
CC	polypeptide is useful in modulating cell-cell interactions of cells
CC	derived from tissues of heart, brain, spinal cord and skeletal muscle.
CC	It is useful in treating and diagnosing chondro sarcoma, atherosclerosis,
CC	Alzheimer's disease, restenosis, ischemic reperfusion, obesity, intimal
CC	hyperplasia and tumors of heart, brain, and spinal cord. The zidm1
CC	polypeptide is also useful in identifying its new family members,
CC	antagonists, agonists and antibodies. Antagonists, antibodies and fusion
CC	proteins of zidm1 are useful in inhibiting platelet aggregation,
CC	apoptosis, neurogenesis and myogenesis. Agonists and antagonists are
CC	useful in studying cell-cell interactions, arthritis, myogenesis,
CC	neurogenesis, connective tissue disorders, chondrogenesis, tumour
CC	proliferation and suppression, extracellular matrix proteins, repair
CC	and remodelling of ischemic reperfusion, inflammation, and apoptosis.
XX	
SO	Sequence 2088 BP; 412 A; 178 C; 381 G; 262 T; 855 other;
Alignment Scores:	
Pred. No.:	1.63e-24
Score:	285.00
Percent Similarity:	81.82%
Best Local Similarity:	81.82%
Query Match:	85.84%
DB:	21
US-09-634-252A-4_COPY_532_586 (1-55) x_AA245827 (1-2088)	
OY	1 CysSerLeuSerArgGlyAlaHisCysSerArgGlyProCysCysAsnAsnThrSerCys 20
Db	1225 TGVSNTNMTNMNAAYGGNGCNCAITGTYSNGAYGAGNCNCCTGTGTAAVAALACMWSNTGY 1286
OY	21 LeuPheGlnProArGgLYTYTGtLcYsaRgaApAlaValaEngUlcYsaSPDIEthrLU 40
Db	1285 YTNHTTCARCNCMGNGTATGARTGYWGNGAYGCNGTNAAYGARTGGAAYATHACNGAR 1344
OY	41 TYrCYstHrgLYaaPserGtYLcInCYsePrOPraSnleuHISLVs 55

Db 1345 TAYTGACNGNGAYWSNGNCARTGYCCNCCNAAYTNCAYAR 1389

Search completed: October 21, 2003, 10:34:17  
Job time : 148.862 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:21:31 ; Search time 1304.78 Seconds  
(without alignments)  
1024.499 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGACSDGPCNNNTSC.....CDITEYCTGDSGQCPNLMHK 55

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 1215238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09634252/unat.21102003.090324.3331/app.query.fasta.1.462  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human10.cdi -LIST=45  
-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09634252 @CGN 1.1 3596 @runat.21102003.090324.3331 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_ylt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rhg:\*  
26: em\_gss\_vhl:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	553	10	AW961362 ESTJ73434
2	332	100.0	728	13	BUS39967 AGENCOURT
3	332	100.0	752	10	BF691008 B62247043
4	328	98.8	340	14	F08148 HSC2RF011.n
5	320	96.4	722	12	BI184544 UNL-P-FN-
6	319	96.1	470	14	CB733322 AMGNNUC:N
7	319	96.1	2891	11	AK046677 Mus muscu
8	319	96.1	3433	11	AK029301 Mus muscu
9	319	96.1	3498	11	AK034022 Mus muscu
10	318	95.8	432	12	BI018895 IL3-MT026
11	308	92.8	261	9	AA317222 EST19173
12	305	91.9	652	9	AV340595 AV340595
13	299	90.1	402	12	BI019679 BI019679
14	293	88.3	326	14	RI15038 RI15038
15	286.5	86.3	648	13	BUS396219 BUS396219
16	286.5	86.3	660	13	BUS395688 BUS395688
17	286.5	86.3	661	13	BUS389998 BUS389998
18	271.5	81.8	579	12	BI844587 BI844587
19	271.5	81.8	583	12	BM186739 BM186739
20	269	81.0	405	14	R52569 R52569
21	268	80.7	428	9	AA718688 AA718688
22	266.5	80.3	664	12	BI429326 BI429326
23	264.5	79.7	598	12	BI981504 BI981504
24	264.5	79.7	616	12	BI476560 BI476560
25	264.5	79.7	652	12	BI429301 BI429301
26	262.5	79.1	648	13	BUS31413 BUS31413
27	260.5	78.5	599	12	BM070858 BM070858
28	259	78.0	447	10	BF905823 BF905823
29	226.5	68.2	512	12	BI839501 BI839501
30	213	64.2	443	14	CB788730 CB788730
31	213	64.2	513	14	CB716849 CB716849
32	213	64.2	609	14	CB580769 CB580769
33	213	64.2	704	14	BY722939 BY722939
34	212	63.9	471	12	CB732873 CB732873
35	210	63.3	526	12	BI839959 BI839959
36	207	62.3	570	13	BQ833435 BQ833435
37	201	60.5	481	13	BY255403 BY255403
38	190	57.2	375	9	AA511039 AA511039
39	185.5	55.9	397	10	AM898896 AM898896
40	185	55.7	690	12	BJ059956 BJ059956
41	184	55.4	618	12	BJ495352 BJ495352
42	176	53.0	622	14	CB578552 CB578552
43	171	51.5	400	14	CB523459 CB523459
44	170	51.2	583	12	BI032916 BI032916
45	168.5	50.8	334	9	AL047647 AL047647

ALIGNMENTS

RESULT 1  
LOCUS AW961362  
DEFINITION ESTJ73434 MAGC resequences, MAGC Homo sapiens CDNA, mRNA sequence.  
ACCESSION AW961362  
VERSION AW961362.1 GI:8151046  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 553)  
Hegde,P., Qi,R., Abernathy,K., Dharep,S., Gaspard,R., Gay,C., Holt

JOURNAL  
TITLE  
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
COMMENT  
Unpublished  
Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnqu@igr.org  
Plate: 157  
Seq primer: Reverse.  
Location/Qualifiers  
1..553  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1ib="MAGE ressequencing, MAGC"  
/note="Vector: pBluescriptSKm"

BASE COUNT 149 a 113 c 148 g 143 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9,48e-28 Length: 553  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x AW961362 (1-553)

QY 1 CysSerLeuSerArgGlyAlaHisCysSerArgGlyProCysCysAsnAsnThrSerCys 20  
47 TGTTCCTCTCCACAGGGGCTCAGTCAGCAGCGGGCTCTGTACCAATCCTCATGT 106  
DB 21 LeuphegInProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
107 CTTTTCAGCCACGAGGGGTATGATGATGCGGGATGCTGTACGAGTGATATTCTGAA 166  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55  
167 TATTGTACTGAGACTGTGTCAGTCCACCAAAATCTTCATTAAG 211  
DB

RESULT 2  
LOCUS BUS39967 728 bp mRNA linear EST 13-SEP-2002  
DEFINITION AGENCOURT\_10254064 NIH\_MGC\_128 Homo sapiens cDNA clone  
IMAGE:6570971 5', mRNA sequence.  
ACCESSION BUS39967  
VERSION BUS39967.1 GI:22850408  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
1 (bases 1 to 728)  
NIH-MGC http://mgc.mci.nih.gov/  
Unpublished  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: NCI  
cDNA Library Preparation: Michael Brownstein Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LNCM2760 row: 1 column: 11  
High quality sequence stop: 527.  
Location/Qualifiers

FEATURES

source  
1..728  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1ib="MAGE:6570971"  
/issue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_1ib="NIH\_MGC\_128"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcgtggc);  
Site 2: SfiI (ggccgtcgtggc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
used in cloning as follows:  
5'-AAGCAGTGTGATCAACGAGAGTGGCCATTAAGCCCGGG-3' and  
5'-ATTCTAGAGCCGAGCGCCGACATG-dT(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the >2 kb  
size fraction (other fractions present in NIH\_MGC\_126 and  
NIH\_MGC\_127). Library created in the laboratory of T.  
Uedlin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC  
Library."

BASE COUNT 213 a 142 c 210 g 162 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.36e-27 Length: 728  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x BUS39967 (1-728)

QY 1 CysSerLeuSerArgGlyAlaHisCysSerArgGlyProCysCysAsnAsnThrSerCys 20  
365 TGTTCCTCTCCACAGGGGCTCAGTCAGCAGCGGGCTCTGTACCAATCCTCATGT 424  
DB 21 LeuphegInProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40  
425 CTTTTCAGCCACGAGGGGTATGATGATGCGGGATGCTGTACGAGTGATATTCTGAA 484  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55  
485 TATTGTACTGAGACTGTGTCAGTCCACCAAAATCTTCATTAAG 529  
DB

RESULT 3  
LOCUS BF691008 752 bp mRNA linear EST 22-DEC-2000  
DEFINITION 602247043F1 NIH\_MGC\_62 Homo sapiens cDNA clone IMAGE:4332275 5',  
mRNA sequence.  
ACCESSION BF691008  
VERSION BF691008.1 GI:11976416  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
1 (bases 1 to 752)  
NIH-MGC http://mgc.mci.nih.gov/  
Unpublished  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:  
http://img1.lnl.gov  
Plate: LNCM198 row: e column: 12  
High quality sequence stop: 719.  
Location/Qualifiers

## FEATURES

source

1..752

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4332275"  
/tissue\_type="melanotic melanoma, high MDR"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_1ib="NIH MG6.62"  
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggccgctcggcc); Site\_2: SfiI (ggccatagggc);  
SfiI (ggccgctcggcc); Site 2: SfiI (ggccatagggc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCATTATGCG-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

BASE COUNT 195 a 151 c 217 g 189 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,42e-27 Length: 752  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x BF691008 (1-752)

OY 1 CysSerLeuSerAenGlyAlaHisCysSerAenGlyProCysCysAsnAsnThrSerCys 20  
Db 73 TGTTCCTCTCCACAGGGGCTCAGCGAGCGAGCGGCTGTGTAACAATACCTCATGT 132  
OY 21 LeupheginProAenGlyTYrGluCysArgAspAlaValaAsnGluCysAspIleThrGlu 40  
Db 133 CTTTTCAGCCACGAGGATGAATGCCGGATGCTGTGAACGAGTGATATTACTGAA 192  
OY 41 TYrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
Db 193 TATTGTACTGAGACTGTGTCAGTCCACCAATCTTCAATAG 237

## RESULT 4

F08148 340 bp mRNA linear EST 21-FEB-1995

LOCUS HSCRF011 normalized infant brain cDNA Homo sapiens cDNA clone  
DEFINITION F08148  
F08148  
F08148.1 GI:677656

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 340)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Auffray, C., Behar, G., Bois, F., Bouchier, C., de Silva, C., Devignes,  
'M.D., Duprat, S., Houligatte, R., Jumeau, M.N., Iamy, B., Lorenzo, F.,  
Mitchell, H., Mariage-Samson, R., Pietu, G., Poullet, Y.,  
Sebastien-Kabakchis, C. and Tessier, A.  
IMAGE: molecular integration of the analysis of the human genome  
and its expression  
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
95277534  
7757816  
Contact: Genethon  
Genexpress-Genethon

Genethon Centre de recherche sur le Genome Humain  
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE  
Tel: 33169472800  
Fax: 33160778698  
Email: genexpress@genethon.fr  
Single read.  
Genexpress library\_id: C; Genexpress\_sequence\_id: yic-2rf01  
Seq primer: (-21)M3 universal.  
Location/Qualifiers

## FEATURES

source

1..340

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="C-2rf01"  
/sex="Female"  
/tissue\_type="total brain"  
/dev\_stage="3 months old"  
/clone\_1ib="normalized infant brain cDNA"  
/note="Organ: brain; Vector: latmid BA; Site\_1: HindIII;  
Site\_2: NotI; sex:Female; dev stage=3 months old;  
isolate=muscular atrophy patient; tissue type=total brain  
; total mRNA was oligo-(dt) primed and directionally  
cloned 5' -> 3' into the HindIII -> NotI sites of the  
latmid BA vector. Clone library from B. Soares, Psychiatry  
Dept. Columbia University, USA. Normalization method:  
Bento Soares, P.N.A.S. in press"

BASE COUNT 87 a 70 c 90 g 84 t 9 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1,41e-27 Length: 340  
Score: 328.00 Matches: 54  
Percent Similarity: 98.18% Conservative: 0  
Best Local Similarity: 98.18% Mismatches: 1  
Query Match: 98.80% Indels: 0  
DB: 14 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x F08148 (1-340)

OY 1 CysSerLeuSerAenGlyAlaHisCysSerAenGlyProCysCysAsnAsnThrSerCys 20  
Db 132 TGTTCCTCTCCACAGGGGCTCAGCGAGCGAGCGGCTGTGTAACAATACCTCATGT 191  
OY 21 LeupheginProAenGlyTYrGluCysArgAspAlaValaAsnGluCysAspIleThrGlu 40  
Db 192 CTTTTCAGCCACGAGGATGAATGCCGGATGCTGTGAACGAGTGATATTACTGAA 251  
OY 41 TYrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
Db 252 TATTGTACTGAGACTGTGTCAGTCCACCAATCTTCAATAG 296

## RESULT 5

B1184544/c 722 bp mRNA linear EST 10-JUL-2001

LOCUS B1184544  
DEFINITION UNL-P-FN-co-e-05-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone  
UNL-P-FN-co-e-05-0-UNL.3', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

EST.  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
1 (bases 1 to 722)  
Caetano, A.R., Johnson, R.K. and Pomp, D.  
Genethon and sequence characterization of a normalized cDNA  
library from swine ovarian follicles  
Mamm. Genome 14 (1), 65-70 (2003)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Pomp, D  
Department of Animal Science  
University of Nebraska, Lincoln



TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Maitahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuno, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Akakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kankawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleisemann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L.M., Staahl, F., Suzuki, F., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojling, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gietzung, S., Hill, D., Hofmann, M., Hume, D.A., Kamita, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, D., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kontutski, S. and Hayashizaki, Y.  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akiyama, T., Akakawa, T., Eono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirose, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
GENOME  
FEATURES  
SOURCE

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.  
Location/Qualifiers  
1..2891  
/organism="Mus musculus"  
/mol\_type="mRNA"

misc\_feature  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:B430307120"  
/db\_xref="taxon:10090"  
/clone="B430307120"  
/sex="male"  
/tissue\_type="adipose"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="4 days neonate"  
1..2891  
/note="a disintegrin and metalloprotease domain 23 (MGI:1345162, GB|NM\_011780, evidence: BLASTN, 100%, match=2431)"

BASE COUNT 721 a 794 c 778 g 598 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,48e-25 Length: 2891  
Score: 319.00 Matches: 53  
Percent Similarity: 96.36% Conservative: 0  
Best Local Similarity: 96.36% Mismatches: 2  
Query Match: 96.08% Indels: 0  
DB: 11 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x AK046677 (1-2891)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20  
Db 1906 TGTTGGCTCTCCAAATGGGGCCCACTGCAGTGCAGGGCCCTGCTGTAAACAACACCTCATGT 1965

QY 21 LeuPheGlnProArgGlyTyrGlyCysArgAspAlaValAsnGlyCysAspIleThrGlu 40  
Db 1966 CTTTTCAGTCACGAGGAGTGAATGTCGGAGTGCCTAAACAGCTGTGATATCACCAG 2025

QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIys 55  
Db 2026 TACTGCACGTGAGACTCTGGCCAGTGCACCGAACCCTCATATAA 2070

RESULT 8  
AK029301  
LOCUS  
DEFINITION  
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4832416K07 product:a disintegrin and metalloprotease domain 23, full insert sequence.  
ACCESSION  
AK029301  
VERSION  
AK029301.1 GI:26081272  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Maitahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuno, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Mateno, Y., Nikaio, I., Pesole, G.,  
Quekembush, U., Schriml, L. M., Staudil, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazza, R. I., Mombere, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sakai, H.,  
Sato, K., Schonbach, C., Seiya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Tayo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kottucki, S.,  
and Hayashizaki, Y.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
5

The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
AUTHORS  
6 (bases 1 to 3433)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kawaga, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ono, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shikaki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M., and Hayashizaki, Y.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan [E-mail: genome-research@riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216]

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/

FEATURES  
SOURCE

1.3433  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4832416K07"  
/db\_xref="taxon:10090"  
/clone="4832416K07"  
/cissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"

misc\_feature  
1.3433  
/note="a disintegrin and metalloprotease domain 23  
(MED|MGI:1345162, GBLNM\_011780, evidence: BLASTN, 100%,  
match=2431)"  
BASE COUNT 944 a 792 c 836 g 861 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,11e-25 Length: 3433  
Score: 319.00 Matches: 53  
Percent Similarity: 96.36% Conserved: 0  
Best Local Similarity: 96.36% Mismatches: 2  
Query Match: 96.08% Indels: 0  
DB: 11 Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x AK029301 (1-3433)  
QY 1 CysSerLeuSerArgGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20  
DB 1207 TGTTCGCTCTCCATGGGGCCCACTGCGAGTGCAGCGCCCTGCTGAACACACCTCATGT 1266  
QY 21 LeuPheGlnProAlaGlyTyrGluCysArgAspAlaValaIaengluCysAspIleThrIu 40  
DB 1267 CTTTTCAGTCACGAGGATATGAATGTCGGGATCCCGTAAACAGCTGTGATATCACCGAG 1326  
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55  
DB 1327 TACTGCACTGAGAGACTCTGGCCAGTCCCACTGCACTTCCATTA 1371  
RESULT 9  
AK034022  
LOCUS  
DEFINITION  
MUS musculus adult male dienecephalon cDNA, RIKEN full-length  
enriched library, clone:930142P22 product: a disintegrin and  
metalloprotease domain 23, full insert sequence.  
ACCESSION  
AK034022  
VERSION  
AK034022.1 GI:26329622  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
2  
Meth. Enzymol. 303, 19-44 (1999)  
TITLE  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
3  
Normalizaton and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
TITLE  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4  
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, U., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M.,  
Sunt, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishigaki, K.,  
Fujitake, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Matshiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Katsukawa, T., Saito, R., Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischman, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Resple, G., Suzuki, F., Tomita, M., Quackenbush, J., Schrim, L.M., Staub, J., Suzuki, F., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Carninci, P., Barsh, G., Blake, J., Boffelli, D., Bojunga, A., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, U., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyszewski, B., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohno, S., and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660  
11217851

**TITLE**  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

**TITLE**  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 3498)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Eono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kondo, M., Koyas, K., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-42 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-3222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
location/Qualifiers

1. 3498  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM,DB:9330142F22"  
/db\_xref="taxon,10090"  
/clone="9330142F22"  
/sex="male"  
/tissue\_type="diencephalon"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
303\_2375  
/note="unlabeled protein product, a disintegrin and metalloprotease domain 23 (MGDI) [GI:1345162, GB|NM\_011780, evidence: BLASTN, 100%, match=2431]"

**CDS**

/codon\_start=1  
/protein\_id="BAC28550.1"  
/db\_xref="GI:26329623"  
/translatability="MKPGLISRRSLTGTGCSLGTGCGPCSPGPEVAPAPCRLL  
VILLALPALTSSRRPARGAASAPPHMNERAETTLVLADENTLQONSSRNTSYS  
AVKEITLPSRLVYINODESPYHVDLFAHQKNAVHLAOSFJLEAGSKRT  
LDLTLNGLNGLSQYVEIHEDGQKMSKGEHCYHGSIRGVSDPVALSTCNGLHM  
FEDDTFVYMEPLELTDDEKSTGRPHIIOQTLAGOYKQKMLSTDSQDMLPELO  
MLRRRAVNSRSGVEFEEMKYLEMLIVNHDHRTYKKRSRSHATNNPAKSVNLDSTY  
KQOLNRVIVAVETWETKDHIDITINPOMLHDFSKYRORIKOHADAVLISRVPH  
YKRSISYREGVCSIRGCVNTEYCPMAVAQVLSOSLQNLQIGIOPSSRKRCET  
ESWGGITMETEYVSHSRFSKCSILEYDFAHQKNAVHLAOSFJLEAGSKRT  
AGEBCDQGVFEVCEYVCKKCSKSNAGHCDGSCNNSTCLFQSRGYECDAVNSCDI  
TEYCTDSGCPNLHKODYSNONGRCNDECKTRDQCQYIGTKAAGSDKFCY  
EKLNTGTEKNGCNGKXGDNRNIPCKSDVFCGFLCTNLTRAPRIQGLQGEIIPISFVH  
QGRVDCR"

**BASE COUNT**  
ORIGIN  
906 a 880 c 853 g 859 t

**Alignment Scores:**  
Pred. No.: 3,196-25 Length: 3498  
Score: 319.00 Matches: 53  
Percent Similarity: 96.36% Conservative: 0  
Best Local Similarity: 96.36% Mismatches: 2  
Query Match: 96.08% Indels: 0  
DB: 11 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x AK034022 (1-3498)

**Qy**  
1 CysSerLeuSerAnglyAlaHisCysSerAspGlyProCysCysAenAenHisSerCys 20  
Db 1887 TGTTCGCTTCATGAGGGCCCACTGAGTACGCGCCCTGTGTATACACACTCATGT 1946

**Qy**  
21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValenGluCysApplleThGlu 40  
Db 1947 CTTTTCAGTCACGAGGATGATGATGCGGATGCCCTTAACACACTGTGATATCACCGAG 2006

**Qy**  
41 TyrCysThrGlyAspSerGlyGlnCysProProAenLeuHisLys 55  
Db 2007 TACTGCACTGAGACACTTGCGCCAGTCCCACTGCACTCATTA 2051

**RESULT 10**  
BI018895 432 bp mRNA linear EST 14-JUN-2001  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI018895 432 bp mRNA linear EST 14-JUN-2001  
H13-MT0267-050101-436-E07 MT0267 Homo sapiens cDNA, mRNA sequence.  
BI018895.1 GI:14425525  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 432)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W., Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Bata, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663  
10737800

**TITLE**  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Genome

Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL3&e2=IL3-MT0267-050101-436-E07&e3=2001-01-05&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 429.  
Location/Qualifiers

# FEATURES

## Source

1. .432  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="MT0267"  
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 119 a 100 c 113 g 100 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.61e-26 Length: 432  
Score: 318.00 Matches: 53  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.15% Mismatches: 0  
Query Match: 95.78% Indels: 0  
DB: 12 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x B101895 (1-432)

OY 2 SerLeuSerAenGlyAlaHisCySerSerApgIyProCyCyAAsnAnThSerCysLeu 21  
|||:|||||  
Db 5 TCCCTCTCTCAGGGGCTCAGCGAGCGAGCCCTGCTGTACATACATCTCATGTCTT 64  
OY 22 PheGlnProArgGlyTyrGlnCysArgAaPaLaValaAenGlyCysAspIleThrGlnTyr 41  
|||:|||||  
Db 65 TTTGAGCCACGAGGGGTATGATGCGGGATGCTGTGAACGAGTGTATATCTGATATAT 124  
OY 42 CysThrGlyAAspSerGlyGlnCysProProAenLeuHisIle 55  
|||:|||||  
Db 125 TGTACTGAGAGACTGTGTGTCAGTCCACCAATCTTCATTAAG 166

Search completed: October 21, 2003, 13:06:58  
Job time : 1308.78 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:34:26 ; Search time 108.654 Seconds  
(without alignments)  
758.947 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599

Perfect score: 630

Sequence: 1 PTKLFEPTECGNGYVEAGEE.....CPPNLHKDQYACNQNOCRC 104

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1840717 seqs, 76133381 residues

Total number of hits satisfying chosen parameters: 3681434

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

MODEL=frame+p2n.model -DEV=xlh  
O=/cgn2\_1/USF70\_sepol/US09634252/runat\_21102003\_090326\_3471/app\_query.fasta\_1.462  
-DB=pending\_patente\_NA\_New -OFMT=faetap -SUFFIX=p2n.rnpn -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bl0sum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=pl0 -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09634252@cgn2\_1.248 @runat\_21102003\_090326\_3471  
-NCPU=6 -ICPU=3 -NO\_MMAP -LANG=QUEBEC -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patente NA New.\*  
2: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	3760	US-60-495-114-329	Sequence 329, App
2	630	100.0	4071	US-60-495-114-330	Sequence 330, App
3	404.5	64.2	1911	US-60-487-610-948	Sequence 948, App
4	404.5	64.2	1911	US-60-485-450-577	Sequence 577, App
5	404.5	64.2	3117	US-60-495-114-825	Sequence 825, App
6	404.5	64.2	3167	US-60-487-610-947	Sequence 947, App
7	404.5	64.2	3167	US-60-485-450-576	Sequence 576, App
8	404.5	64.2	3185	US-60-487-610-949	Sequence 949, App
9	404.5	64.2	3185	US-60-485-450-578	Sequence 578, App
10	404.5	64.2	3421	US-60-495-114-827	Sequence 827, App
11	404.5	64.2	3431	US-60-495-114-826	Sequence 826, App

12	404.5	64.2	3433	US-60-495-114-828	Sequence 828, App
13	363.5	57.7	2636	US-60-490-890-1042	Sequence 1042, App
14	363.5	57.7	2744	US-60-490-890-1038	Sequence 1038, App
15	363.5	57.7	2775	US-60-495-114-168	Sequence 168, App
16	363.5	57.7	2775	US-60-495-114-173	Sequence 173, App
17	363.5	57.7	2865	US-60-495-114-174	Sequence 174, App
18	363.5	57.7	2876	US-60-495-114-166	Sequence 166, App
19	363.5	57.7	3347	US-60-490-890-1040	Sequence 1040, App
20	363.5	57.7	3759	US-60-495-114-172	Sequence 172, App
21	363.5	57.7	3777	US-60-495-114-169	Sequence 169, App
22	363.5	57.7	3777	US-60-495-114-175	Sequence 175, App
23	363.5	57.7	3801	US-60-495-114-171	Sequence 171, App
24	363.5	57.7	3819	US-60-495-114-167	Sequence 167, App
25	363.5	57.7	3819	US-60-495-114-170	Sequence 170, App
26	363.5	57.7	3867	US-60-495-114-176	Sequence 176, App
27	363.5	57.7	3878	US-60-495-114-165	Sequence 165, App
28	292	46.3	2523	PCT-US03-30720-591	Sequence 591, App
29	292	46.3	3271	US-10-670-184-1	Sequence 1, App1
30	292	46.3	3390	US-10-670-184-2	Sequence 2, App1
31	292	46.3	3445	US-60-495-114-964	Sequence 964, App
32	292	46.3	3482	PCT-US03-30720-267	Sequence 267, App
33	292	46.3	3523	US-60-495-114-965	Sequence 965, App
34	292	46.3	3554	PCT-US03-28227-1868	Sequence 1868, App
35	292	46.3	3582	US-10-670-184-3	Sequence 3, App1
36	292	46.3	3606	PCT-US03-28227-1890	Sequence 1890, App
37	292	46.3	3721	PCT-US03-28227-1889	Sequence 1889, App
38	278	44.1	2815	PCT-US03-28251-6	Sequence 6, App1
39	274	43.5	1194	PCT-US03-26780-217	Sequence 217, App
40	274	43.5	2286	US-60-495-114-447	Sequence 447, App
41	274	43.5	2313	US-60-495-114-448	Sequence 448, App
42	274	43.5	2336	US-60-495-114-449	Sequence 449, App
43	274	43.5	2712	US-60-487-610-511	Sequence 511, App
44	274	43.5	2740	PCT-US03-28251-3	Sequence 3, App1
45	274	43.5	2740	US-60-490-890-1712	Sequence 1712, App

#### ALIGNMENTS

RESULT 1  
US-60-495-114-329  
; Sequence 329, Application US/60495114  
; GENERAL INFORMATION:  
; APPLICANT: CARILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CU001480  
; CURRENT APPLICATION NUMBER: US/60/495,114  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 329  
; LENGTH: 3760  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-495-114-329

#### Alignment Scores:

Pred. No.: 7.68e-71 Length: 3760  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-60-495-114-329 (1-3760)

Qy 1 ProthrlyseuplegiupprothrgiucyselysnglytyrValaIgluaIgluGlu 20  
Db 1709 CCAACAAAGCTATTTCAGCCACCGAATGTGAAATGATACGTGGAAGCTGGGAGAG 1768  
Qy 21 CysaspCysgllyphen:valaIgluCytyrGlyleuCyCyslylylyCysSerLeuSer 40

```
Db 1769 TGTGATTGTTTCAATGTCGAATGATTATGCTGTGAAGAAATGTTCCCTCTCC 1828
Qy 41 AenglyAlaHisCysSeraspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60
Db 1829 AACGGGGCTCAGTGCAGCCACGGGGCTCTGTACAAATACCTCATGCTCTTTTTCAGCCA 1888
Qy 61 ArgGlyTYrGluCysArgAspAlaValAengluCysAspIleThrGluTYrCysThrGly 80
Db 1889 CGAGGGTATGAATGCGGGATGCTGTGAACGAGTGTGATATTACTGAATATTGTACTGGA 1948
Qy 81 AaspserGlyGlnCysProProAsnLeuHisIlySGlnAspGlyTYrAlaCysAsnGlnAsn 100
Db 1949 GACTCTGGTCAGTGCAGCCCAAAATCTTCAATAGCAAGACGAGATATGCAATCAAAAT 2008
Qy 101 GlnGlyArgCys 104
Db 2009 CAGGGCGCTGC 2020

RESULT 2
US-60-495-114-330
; Sequence 330, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 4071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-495-114-330

Alignment Scores:
Pred. No.: 8, 61e-71 Length: 4071
Score: 630.00 Matches: 104
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-495-114-330 (1-4071)
Qy 1 ProThrIlyLeuPheGlnProThrGluCysGlyAenglyTYrValGlnAaglGluGlu 20
Db 1709 CCAACAAAGCTATTGAGCCCAAGGAATGTGAATGATGACGTGGAACCTGGGGAGAG 1768
Qy 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysCysIlySlysCysSerLeuSer 40
Db 1769 TGTGATTGTTTCAATGTCGAATGATTATGCTGTGAAGAAATGTTCCCTCTCC 1828
Qy 41 AenglyAlaHisCysSeraspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60
Db 1829 AACGGGGCTCAGTGCAGCCACGGGGCTCTGTACAAATACCTCATGCTCTTTTTCAGCCA 1888
Qy 61 ArgGlyTYrGluCysArgAspAlaValAengluCysAspIleThrGluTYrCysThrGly 80
Db 1889 CGAGGGTATGAATGCGGGATGCTGTGAACGAGTGTGATATTACTGAATATTGTACTGGA 1948
Qy 81 AaspserGlyGlnCysProProAsnLeuHisIlySGlnAspGlyTYrAlaCysAsnGlnAsn 100
Db 1949 GACTCTGGTCAGTGCAGCCCAAAATCTTCAATAGCAAGACGAGATATGCAATCAAAAT 2008
Qy 101 GlnGlyArgCys 104
Db 2009 CAGGGCGCTGC 2020

RESULT 3
US-60-487-610-948
```

```
; Sequence 948, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-948

Alignment Scores:
Pred. No.: 4, 83e-42 Length: 1911
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-487-610-948 (1-1911)
Qy 1 ProThrIlyLeuPheGlnProThrGluCysGlyAenglyTYrValGlnAaglGluGlu 20
Db 1312 CCCCTCAAGCTCTGAGCCCCCAAGATGCGGGAAGCGCTTGTGAGAGCGAGGGAGAG 1371
Qy 21 CysAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysIlySlysCys 37
Db 1372 TCGCATCGCGCTCGCTCGTCAGAGATGCAGCCGCGAGGTGGCAATGCTGCAGAAATGC 1431
Qy 38 SerLeuSerAenglyAlaHisCysSeraspGlyProCysCysAsnAsnThrSerCysLeu 57
Db 1432 ACCCTGACTCAGCAGCGCATGTGACGACGCGGCTCTGCTGCGCCG-----TGCAG 1485
Qy 58 PheGlnProAspArgGlyTYrGluCysArgAspAlaValAengluCysAspIleThrGluTYr 77
Db 1486 TACGAACCAACGGGGTGTCTCTGCGGAGAGCGCGTGAACGAGTGTGACATGCGGAGACC 1545
Qy 78 CysThrGlyAaspserGlyGlnCysProProAsnLeuHisIlySGlnAspGlyTYrAlaCys 97
Db 1546 TGCACCGGGGACTTACGACAGTGCAGCCGCTTAACCTGCAGCAAGCTGAGAGTTACTACTGT 1605
Qy 98 AengluAengluGlyArgCys 104
Db 1606 GACCATGTAGCAGGGCGCTGC 1626

RESULT 4
US-60-485-450-577
; Sequence 577, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-577
```

Alignment Scores:  
Pred. No.: 4,83e-42 Length: 1911  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 7 Gaps: 2

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-60-485-450-577 (1-1911)

OY 1 ProthrllysleuPhehgluprothrglucysgllyanglytyrvalaigluaglygluclu 20  
Db 1312 CCCCTCAAGCTCTGGACCCCAAGAGTGGCGGAACGGCTTCGTGAGGCGGGGAGAG 1371  
OY 21 CysAAspCysglYrhenisvalaigluCysTyr-----GlyLeuCysCysLylysCys 37  
Db 1372 TGGACCTGGCGGTGGGTGCGAGAGTGCAGCGCCGCGAGTGGCAATGCTGCAGAAATGTC 1411  
OY 38 SerleuSeranaglYalanisCysSerAAspGlyProCysCysAAsnAnthrSerCysleu 57  
Db 1432 ACCCTGACTCAGACCGCATGTGCAGCGACGCGCTGTCTGCGACG-----TGCAG 1485  
OY 58 PhehglProAAspGlyTYrGluCysArgAAspAlaValasnGluCysAAspIlethrgluTyr 77  
Db 1486 TACCAACACCGGGGTGTGTCTCTGCGAGAGCGCGTGAACGAGTGCAGATCGCGAGACC 1545  
OY 78 CysAthrGlyAAspSerGlyGlnCysProProAAsnleuHisLySglAAsnGlyTYrAlaCys 97  
Db 1546 TGCACCGGGGACTCTAGCCAGTACCGCTTAACCTGCAGAAAGCTGAGCGTTACTACTGT 1605  
OY 98 AsnGlnAsnGlnGlyArgCys 104  
Db 1606 GACCATGACGACGCGCGCTGC 1626

RESULT 5  
US-60-495-114-825  
Sequence 825, Application US/60495114  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
TITLE OF INVENTION: USES THEREOF  
FILE REFERENCE: CU001480  
CURRENT APPLICATION NUMBER: US/60/495,114  
CURRENT FILING DATE: 2003-08-15  
NUMBER OF SEQ ID NOS: 91238  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 825  
LENGTH: 3117  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-495-114-825

Alignment Scores:  
Pred. No.: 9.78e-42 Length: 3117  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 7 Gaps: 2

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-60-495-114-825 (1-3117)

OY 1 ProthrllysleuPhehgluprothrglucysgllyanglytyrvalaigluaglygluclu 20  
Db 998 CCCCTCAAGCTCTGGACCCCAAGAGTGGCGGAACGGCTTCGTGAGGCGGGGAGAG 1057  
OY 21 CysAAspCysglYrhenisvalaigluCysTyr-----GlyLeuCysCysLylysCys 37  
Db 1058 TGGACCTGGCGGTGGGTGCGAGAGTGCAGCGCCGCGAGTGGCAATGCTGCAGAAATGTC 1117  
OY 38 SerleuSeranaglYalanisCysSerAAspGlyProCysCysAAsnAnthrSerCysleu 57  
Db 1606 GACCATGACGACGCGCGCTGC 1626

Db 1118 ACCCTGACTCAGACCGCATGTGTCAGGACGCGCTGTCTGCGCCG-----TGCAG 1171  
OY 58 PhehglProAAspGlyTYrGluCysArgAAspAlaValasnGluCysAAspIlethrgluTyr 77  
Db 1172 TACCAACACCGGGGTGTGTCTCTGCGAGAGCGCTGAACAGAGTGCAGATCGCGAGACC 1231  
OY 78 CysAthrGlyAAspSerGlyGlnCysProProAAsnleuHisLySglAAsnGlyTYrAlaCys 97  
Db 1232 TGCACCGGGGACTCTAGCCAGTACCGCTTAACCTGCAGAAAGCTGAGCGTTACTACTGT 1291

OY 98 AsnGlnAsnGlnGlyArgCys 104  
Db 1292 GACCATGACGACGCGCGCTGC 1312

RESULT 6  
US-60-487-610-947  
Sequence 947, Application US/60487610  
GENERAL INFORMATION:  
APPLICANT: HUANG, Hongjin  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS.  
TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CU001469  
CURRENT APPLICATION NUMBER: US/60/487,610  
CURRENT FILING DATE: 2003-07-17  
NUMBER OF SEQ ID NOS: 97101  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 947  
LENGTH: 3167  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-487-610-947

Alignment Scores:  
Pred. No.: 1e-41 Length: 3167  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 7 Gaps: 2

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-60-487-610-947 (1-3167)

OY 1 ProthrllysleuPhehgluprothrglucysgllyanglytyrvalaigluaglygluclu 20  
Db 1312 CCCCTCAAGCTCTGGACCCCAAGAGTGGCGGAACGGCTTCGTGAGGCGGGGAGAG 1371  
OY 21 CysAAspCysglYrhenisvalaigluCysTyr-----GlyLeuCysCysLylysCys 37  
Db 1372 TGGACCTGGCGGTGGGTGCGAGAGTGCAGCGCCGCGAGTGGCAATGCTGCAGAAATGTC 1431  
OY 38 SerleuSeranaglYalanisCysSerAAspGlyProCysCysAAsnAnthrSerCysleu 57  
Db 1432 ACCCTGACTCAGACCGCATGTGCAGCGACGCGCTGTCTGCGCCG-----TGCAG 1485  
OY 58 PhehglProAAspGlyTYrGluCysArgAAspAlaValasnGluCysAAspIlethrgluTyr 77  
Db 1486 TACCAACACCGGGGTGTGTCTCTGCGAGAGCGCGTGAACGAGTGCAGATCGCGAGACC 1545  
OY 78 CysAthrGlyAAspSerGlyGlnCysProProAAsnleuHisLySglAAsnGlyTYrAlaCys 97  
Db 1546 TGCACCGGGGACTCTAGCCAGTACCGCTTAACCTGCAGAAAGCTGAGCGTTACTACTGT 1605  
OY 98 AsnGlnAsnGlnGlyArgCys 104  
Db 1606 GACCATGACGACGCGCGCTGC 1626

RESULT 7  
US-60-485-450-576  
Sequence 576, Application US/60485450  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele

```

; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INJECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 576
; LENGTH: 3167
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-576

Alignment Scores:
Pred. No.: 1e-41 Length: 3167
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 7 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-485-450-576 (1-3167)
QY 1 ProthrlybLeuPheGluProThrgluCysGlyAsnGlyTYrValAGluAGluGlu 20
DB 1312 CCCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAGAG 1371
QY 21 CysAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysLysLysCys 37
DB 1372 TCCGACTCGGCTCGGTGTCAGAGAGTGCACCGCGCAGAGTGTGCAACTGTGCAAGAAATGC 1431
QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeu 57
DB 1432 ACCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1485
QY 58 PheGlnProArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrgluTYr 77
DB 1486 TACGACACCAAGCGGCTGTCTCTGCGAGAGCGCGTGAAGAGTGCACATCCGCGAGACC 1545
QY 78 CysThrglyAspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrAlaCys 97
DB 1546 TGCACCGGGAGCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1605
QY 98 AsnGlnAsnGlnGlyArgCys 104
DB 1606 GACCATGAGCAGGCGCGCTGC 1626

RESULT 8
US-60-487-610-949
; Sequence 949, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INJECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 949
; LENGTH: 3185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-949

Alignment Scores:
Pred. No.: 1.01e-41 Length: 3185
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 7 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-487-610-949 (1-3185)
QY 1 ProthrlybLeuPheGluProThrgluCysGlyAsnGlyTYrValAGluAGluGlu 20
DB 1312 CCCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAGAG 1371
QY 21 CysAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysLysLysCys 37
DB 1372 TCCGACTCGGCTCGGTGTCAGAGAGTGCACCGCGCAGAGTGTGCAACTGTGCAAGAAATGC 1431
QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeu 57
DB 1432 ACCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1485
QY 58 PheGlnProArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrgluTYr 77
DB 1486 TACGACACCAAGCGGCTGTCTCTGCGAGAGCGCGTGAAGAGTGCACATCCGCGAGACC 1545
QY 78 CysThrglyAspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrAlaCys 97
DB 1546 TGCACCGGGAGCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1605
QY 98 AsnGlnAsnGlnGlyArgCys 104
DB 1606 GACCATGAGCAGGCGCGCTGC 1626
```

```

Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 7 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-487-610-949 (1-3185)
QY 1 ProthrlybLeuPheGluProThrgluCysGlyAsnGlyTYrValAGluAGluGlu 20
DB 1312 CCCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAGAG 1371
QY 21 CysAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysLysLysCys 37
DB 1372 TCCGACTCGGCTCGGTGTCAGAGAGTGCACCGCGCAGAGTGTGCAACTGTGCAAGAAATGC 1431
QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeu 57
DB 1432 ACCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1485
QY 58 PheGlnProArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrgluTYr 77
DB 1486 TACGACACCAAGCGGCTGTCTCTGCGAGAGCGCGTGAAGAGTGCACATCCGCGAGACC 1545
QY 78 CysThrglyAspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrAlaCys 97
DB 1546 TGCACCGGGAGCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1605
QY 98 AsnGlnAsnGlnGlyArgCys 104
DB 1606 GACCATGAGCAGGCGCGCTGC 1626

RESULT 9
US-60-485-450-578
; Sequence 578, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INJECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CLO01470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 578
; LENGTH: 3185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-578

Alignment Scores:
Pred. No.: 1.01e-41 Length: 3185
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 7 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-485-450-578 (1-3185)
QY 1 ProthrlybLeuPheGluProThrgluCysGlyAsnGlyTYrValAGluAGluGlu 20
DB 1312 CCCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAGAG 1371
QY 21 CysAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysLysLysCys 37
DB 1372 TCCGACTCGGCTCGGTGTCAGAGAGTGCACCGCGCAGAGTGTGCAACTGTGCAAGAAATGC 1431
QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeu 57
DB 1432 ACCCTCAAGCTCTGTGAGACCCCGAGAGCGGCGGAAAGCGCTTGTGAGAGCGAGAG 1485
```

```
OY 58 PheGlnProArgIYTYrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyr 77
Db 1486 TACGACCAACGGGGGTGTCTCTGCGAGAGCCCGTGAACGAGTGCACATCGCGAGACC 1545
OY 78 CysThrGlyAAspSerGlyGlnCysProProAsnLeuHisIleGlnAspGlyTYrAlaCys 97
Db 1546 TGCACCGGGGACTTACGAGTGCAGCCCTTAACCTGCACAGCTGACGGTTACTACTGT 1605
OY 98 AsnGlnAsnGlnGlyArgCys 104
Db 1606 GACCATGACACAGGCCGCTGC 1626

RESULT 10
US-60-495-114-827
: Sequence 827, Application US/60495114
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CLO01480
: CURRENT APPLICATION NUMBER: US/60/495,114
: CURRENT FILING DATE: 2003-08-15
: NUMBER OF SEQ ID NOS: 91238
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 827
: LENGTH: 3421
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-495-114-827

Alignment Scores:
Pred. No.: 1,12e-41 Length: 3421
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 7 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-60-495-114-827 (1-3421)
OY 1 ProThrIleuLeuPheGlnProThrGluCysGlyAsnGlyTYrValGluAlaGlyGluGlu 20
Db 1312 CCCCTCAGAGCTCTCTGACCCCCAGAGTGCAGGAAAGCGCTTCTGTGACGACGAGGAGAG 1371
OY 21 CysAAspCysGlyPheHisValGluCysTYr-----GlyLeuCysCysIleCysCys 37
Db 1372 TGCAGACTCGGCTCGGTGCAGAGTGCAGCCCGCAGGTCGCAACTGCTGCAGAAATGC 1431
OY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnIleThrSerCysIleu 57
Db 1432 ACCCTGACTCAGACGCGCATGTGCAGCGACGCGGCTCTGTCTGCGCGC-----TGCAG 1485
OY 58 PheGlnProArgIYTYrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyr 77
Db 1486 TACGACCAACGGGGGTGTCTCTGCGAGAGCCCGTGAACGAGTGCACATCGCGAGACC 1545
OY 78 CysThrGlyAAspSerGlyGlnCysProProAsnLeuHisIleGlnAspGlyTYrAlaCys 97
Db 1546 TGCACCGGGGACTTACGAGTGCAGCCCTTAACCTGCACAGCTGACGGTTACTACTGT 1605
OY 98 AsnGlnAsnGlnGlyArgCys 104
Db 1606 GACCATGACACAGGCCGCTGC 1626
```

Search completed: October 21, 2003, 14:42:51  
Job time : 216.654 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 12:02:06 ; Search time 264.306 Seconds  
(without alignments)  
1052.855 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599

Perfect score: 630

Sequence: 1 PTKLFEPTCEGNGYVEAGE.....CPPLHKQDGVACNONGRC 104

Scoring table:

BLOSUM62			
Xgapop 10.0 ,	Xgapext 0.5		
Ygapop 10.0 ,	Ygapext 0.5		
Fgapop 6.0 ,	Fgapext 7.0		
Delpop 6.0 ,	Delpext 7.0		

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3534790

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh  
-Q=/cgn2/1/USPRO.spool/US09634252/runat\_21102003\_090327\_3499/epq\_query.fasta\_1.462  
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -MATRIX=b10sum2  
-TRANS=human40.cdi -DIST=45 -DCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09634252 @CNC 1.1 339 @runat\_21102003\_090327\_3499  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSELOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELDPXT=7

Database: Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	2268	9	US-09-809-790-1 Sequence 1, Appl1

2	630	100.0	2268	10	US-09-809-617-1	Sequence 1, Appl1
3	630	100.0	2268	12	US-10-260-506-1	Sequence 1, Appl1
4	630	100.0	2499	14	US-10-202-675-2	Sequence 2, Appl1
5	630	100.0	4043	9	US-09-824-129-2	Sequence 2, Appl1
6	630	100.0	4043	12	US-10-359-464-2	Sequence 2, Appl1
7	630	100.0	4720	12	US-10-260-506-9	Sequence 9, Appl1
8	618	98.1	494	11	US-09-918-995-22101	Sequence 22101, A
9	567	90.0	1668	9	US-09-792-2008-17	Sequence 17, Appl1
10	567	90.0	2088	9	US-09-809-790-3	Sequence 3, Appl1
11	567	90.0	2088	10	US-09-809-617-3	Sequence 3, Appl1
12	567	90.0	2088	12	US-10-260-506-3	Sequence 3, Appl1
13	567	90.0	2127	14	US-10-260-506-11	Sequence 11, Appl1
14	363.5	57.7	2604	14	US-10-202-675-5	Sequence 5, Appl1
15	354.5	56.3	494	11	US-09-918-995-2610	Sequence 2610, Ap
16	338.5	53.7	1674	9	US-09-792-2008-15	Sequence 15, Appl1
17	292	46.3	3431	12	US-10-177-308-1	Sequence 1, Appl1
18	292	46.3	3468	12	US-10-177-308-3	Sequence 3, Appl1
19	292	46.3	3582	12	US-10-199-672-465	Sequence 465, App
20	292	46.3	3582	12	US-10-187-749-465	Sequence 465, App
21	292	46.3	3582	12	US-10-194-457-465	Sequence 465, App
22	292	46.3	3582	12	US-10-184-642-465	Sequence 465, App
23	292	46.3	3582	12	US-10-196-747-465	Sequence 465, App
24	292	46.3	3582	12	US-10-173-689-465	Sequence 465, App
25	292	46.3	3582	12	US-10-173-690-465	Sequence 465, App
26	292	46.3	3582	12	US-10-173-691-465	Sequence 465, App
27	292	46.3	3582	12	US-10-173-692-465	Sequence 465, App
28	292	46.3	3582	12	US-10-173-694-465	Sequence 465, App
29	292	46.3	3582	12	US-10-173-698-465	Sequence 465, App
30	292	46.3	3582	12	US-10-173-699-465	Sequence 465, App
31	292	46.3	3582	12	US-10-173-707-465	Sequence 465, App
32	292	46.3	3582	12	US-10-174-569-465	Sequence 465, App
33	292	46.3	3582	12	US-10-174-583-465	Sequence 465, App
34	292	46.3	3582	12	US-10-174-587-465	Sequence 465, App
35	292	46.3	3582	12	US-10-174-589-465	Sequence 465, App
36	292	46.3	3582	12	US-10-174-591-465	Sequence 465, App
37	292	46.3	3582	12	US-10-175-736-465	Sequence 465, App
38	292	46.3	3582	12	US-10-175-742-465	Sequence 465, App
39	292	46.3	3582	12	US-10-175-744-465	Sequence 465, App
40	292	46.3	3582	12	US-10-175-745-465	Sequence 465, App
41	292	46.3	3582	12	US-10-175-748-465	Sequence 465, App
42	292	46.3	3582	12	US-10-175-751-465	Sequence 465, App
43	292	46.3	3582	12	US-10-175-754-465	Sequence 465, App
44	292	46.3	3582	12	US-10-176-480-465	Sequence 465, App
45	292	46.3	3582	12	US-10-176-489-465	Sequence 465, App

#### ALIGNMENTS

RESULT 1  
US-09-809-790-1  
Sequence 1, Application US/09809790  
Patent No. US20020072102A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baindur, Nand  
APPLICANT: Delsner, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/809,790  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)..  
NAME/KEY: misc\_feature

LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-09-809-790-1

## Alignment Scores:

Pred. No.:	1,456-69	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-809-790-1 (1-2268)

QY 1 ProthrlYsLeuphEgLuProthrgLucYsGlyAengLYTYrVaIGluAIGlyGlu 20  
DB 1119 CCAACAAAGCTATTGGAGCCCAAGAAATGTAATGATGCTGGAAGCTGGAGAGAG 1178  
QY 21 CysaPcySGlyPheHIsVaIGluCYstYrGlyLeuCYsGlyLysCYsSerLeuSer 40  
DB 1179 TGTGATTGCTTTTCATGTCGAATGCTATGATGCTGTAAGAAATGTTCCCTCTCC 1238  
QY 41 AsnglyAIAhIsCYsSerAspGlyProCYsCYsAnaenThrSerCYsLeuphEgLnPro 60  
DB 1239 AACGGGGCTCACTGACAGCAGCGGCCCTGCTTAACAATACCTCAAGTCTTTTCAGCCA 1298  
QY 61 ArgGlyTYrGluCYsArgAspAlaValAsngluCYsAspIleThrGluTYrCYsthrGly 80  
DB 1299 CGAGGTTATGATGCGGGATGCTGTGAACGATGTGATTAAGTAATTTGTAAGTGA 1358  
QY 81 AspSerGlyGlnCYsProProAsnLeuHIsIlysgInaSpGlyTYrAlaCYsAsngInAsn 100  
DB 1359 GACTGTGCTAGTGGCCCAACAATCTTCAATAGCAAGCGATATGCAATCAATCAAAAT 1418  
QY 101 GlnGlyArgCYs 104  
DB 1419 CAGGGCCGCTGC 1430

## RESULT 2

US-09-809-617-1  
Sequence 1, Application US/09809617  
Patent No. US20020137178A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baidur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/809,617  
CURRENT FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: US/09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-09-809-617-1

## Alignment Scores:

Pred. No.:	1,456-69	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 10 Gaps: 0  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-809-617-1 (1-2268)

QY 1 ProthrlYsLeuphEgLuProthrgLucYsGlyAengLYTYrVaIGluAIGlyGlu 20  
DB 1119 CCAACAAAGCTATTGGAGCCCAAGAAATGTAATGATGCTGGAAGCTGGAGAGAG 1178  
QY 21 CysaPcySGlyPheHIsVaIGluCYstYrGlyLeuCYsGlyLysCYsSerLeuSer 40  
DB 1179 TGTGATTGCTTTTCATGTCGAATGCTATGATGCTGTAAGAAATGTTCCCTCTCC 1238  
QY 41 AsnglyAIAhIsCYsSerAspGlyProCYsCYsAnaenThrSerCYsLeuphEgLnPro 60  
DB 1239 AACGGGGCTCACTGACAGCAGCGGCCCTGCTTAACAATACCTCAATGCTTTTCAGCCA 1298  
QY 61 ArgGlyTYrGluCYsArgAspAlaValAsngluCYsAspIleThrGluTYrCYsthrGly 80  
DB 1299 CGAGGTTATGATGCGGGATGCTGTGAACGATGTGATTAAGTAATTTGTAAGTGA 1358  
QY 81 AspSerGlyGlnCYsProProAsnLeuHIsIlysgInaSpGlyTYrAlaCYsAsngInAsn 100  
DB 1359 GACTGTGCTAGTGGCCCAACAATCTTCAATAGCAAGCGATATGCAATCAATCAAAAT 1418  
QY 101 GlnGlyArgCYs 104  
DB 1419 CAGGGCCGCTGC 1430

## RESULT 3

US-10-260-506-1  
Sequence 1, Application US/10260506  
Publication No. US20030153064A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baidur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Taft, David W.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHP  
FILE REFERENCE: 98-29C1  
CURRENT APPLICATION NUMBER: US/10/260,506  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: 60/092,371  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/147, 410  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 09/351,414  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-10-260-506-1

## Alignment Scores:

Pred. No.:	1,456-69	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-10-260-506-1 (1-2268)



Qy 1 ProthrlvsleuphegluprothrgluCysglYAsnGlyTYrValGluAlaGlyGluGlu 20  
Db 1119 CCAACAAAGCTATTGAGCCACCGAATGTGAAATGATACGTGAAAGCTGGGAGAG 1178  
Qy 21 CysAspCysGlyPheHisValGluCysETyrGlyLeuCysCysLySlyB/CysSerLeuSer 40  
Db 1179 TGTGATTGTGGTTTTCATGTGAAATGCTATGATGATTAAGTAAATAATGTTCCCTCTCC 1238  
Qy 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
Db 1239 AACGGGGCTCACTCAGCGAGCGGGCCCTGCTGAACATATACCTCATGCTTTTCAGCCA 1298  
Qy 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGlyTYrCysThrGly 80  
Db 1299 CGAGGGATGATGATCCCGGATGCTGTGAACGAGTGTATATTACGAATATTGACTGGA 1358  
Qy 81 AspSerGlyGlnCysProProAsnLeuHislySglAspGlyTYrAlaCysAsnGlnAsn 100  
Db 1359 GACTCTGGTCAAGTCCACCAAAATCTTCATAGAAGCGGATATGCAATCAAAAT 1418  
Qy 101 GlnGlyArgCys 104  
Db 1419 CAGGGCCGCTGC 1430

## RESULT 4

US-10-202-675-2  
Sequence 2, Application US/10202675  
Publication No. US20030100091A1

## GENERAL INFORMATION:

APPLICANT: Cerretti, Douglas P.  
TITLE OF INVENTION: SVPB3-13 AND SVPB3-17 DNA AND POLYPEPTIDES  
FILE REFERENCE: 03260.0051-00304  
CURRENT FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: US/10/202, 675  
PRIOR FILING DATE: 2000-08-07  
PRIOR APPLICATION NUMBER: US/09/634, 252  
PRIOR FILING DATE: 2000-07-23  
PRIOR APPLICATION NUMBER: 60/074, 310  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 2499  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-202-675-2

## Alignment Scores:

Pred. No.: 1,64e-69 Length: 2499  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-10-202-675-2 (1-2499)

Qy 1 ProthrlvsleuphegluprothrgluCysglYAsnGlyTYrValGluAlaGlyGluGlu 20  
Db 1486 CCAACAAAGCTATTGAGCCACCGAATGTGAAATGATACGTGAAAGCTGGGAGAG 1545  
Qy 21 CysAspCysGlyPheHisValGluCysETyrGlyLeuCysCysLySlyB/CysSerLeuSer 40  
Db 1546 TGTGATTGTGGTTTTCATGTGAAATGCTATGATGATTAAGTAAATAATGTTCCCTCTCC 1605  
Qy 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
Db 1606 AACGGGGCTCACTCAGCGAGCGGGCCCTGCTGAACATATACCTCATGCTTTTCAGCCA 1665  
Qy 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGlyTYrCysThrGly 80  
Db 1666 CGAGGGATGATGATCCCGGATGCTGTGAACGAGTGTATATTACGAATATTGACTGGA 1725  
Qy 81 AspSerGlyGlnCysProProAsnLeuHislySglAspGlyTYrAlaCysAsnGlnAsn 100

Db 1726 GACTCTGGTCAAGTCCACCAAAATCTTCATAGAAGCGGATATGCAATCAAAAT 1785  
Qy 101 GlnGlyArgCys 104  
Db 1786 CAGGGCCGCTGC 1797

## RESULT 5

US-09-824-129-2  
Sequence 2, Application US/09824129  
Patent No. US2002001840A1

## GENERAL INFORMATION:

APPLICANT: Lopez-Otin, Carlos  
APPLICANT: Miguel, Santiago Cal  
APPLICANT: Freije, Jose Maria Perez  
APPLICANT: Garcia, Jose Manuel Lopez  
APPLICANT: Bianchi, Albert Bernard  
APPLICANT: Trail, Pamela  
TITLE OF INVENTION: Methods and Compositions for Modulating  
Integrin-mediated Cell-Cell Interactions  
FILE REFERENCE: D0015-NP  
CURRENT FILING DATE: 2001-04-02  
CURRENT APPLICATION NUMBER: US/09/824, 129  
PRIOR FILING DATE: 2000-04-03  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4043  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-824-129-2

## Alignment Scores:

Pred. No.: 3.11e-69 Length: 4043  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-09-824-129-2 (1-4043)

Qy 1 ProthrlvsleuphegluprothrgluCysglYAsnGlyTYrValGluAlaGlyGluGlu 20  
Db 2502 CCAACAAAGCTATTGAGCCACCGAATGTGAAATGATACGTGAAAGCTGGGAGAG 2561  
Qy 21 CysAspCysGlyPheHisValGluCysETyrGlyLeuCysCysLySlyB/CysSerLeuSer 40  
Db 2562 TGTGATTGTGGTTTTCATGTGAAATGCTATGATGATTAAGTAAATAATGTTCCCTCTCC 2621  
Qy 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
Db 2622 AACGGGGCTCACTCAGCGAGCGGGCCCTGCTGAACATATACCTCATGCTTTTCAGCCA 2681  
Qy 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGlyTYrCysThrGly 80  
Db 2682 CGAGGGATGATGATCCCGGATGCTGTGAACGAGTGTATATTACGAATATTGACTGGA 2741  
Qy 81 AspSerGlyGlnCysProProAsnLeuHislySglAspGlyTYrAlaCysAsnGlnAsn 100  
Db 2742 GACTCTGGTCAAGTCCACCAAAATCTTCATAGAAGCGGATATGCAATCAAAAT 2801  
Qy 101 GlnGlyArgCys 104  
Db 2802 CAGGGCCGCTGC 2813  
RESULT 6  
US-10-359-464-2  
Sequence 2, Application US/10359464  
Publication No. US20030143692A1  
GENERAL INFORMATION:  
APPLICANT: Lopez-Otin, Carlos

```

; APPLICANT: Miguel, Santiago Cal
; APPLICANT: Freije, Jose Maria Perez
; APPLICANT: Garcia, Jose Manuel Lopez
; APPLICANT: Bianchi, Albert Bernard
; APPLICANT: Trail, Pamela
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING INTEGRIN-MEDIATED
; FILE REFERENCE: CELL-CELL INTERACTIONS
; CURRENT APPLICATION NUMBER: US/10/359,464
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 09/824,129
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 4043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-359-464-2

Alignment Scores:
Pred. No.: 3,11e-69 Length: 4043
Score: 630.00 Matches: 104
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-634-252a-4_COPY_496_599 (1-104) x US-10-359-464-2 (1-4043)

QY 1 ProthrLysLeuPheGluProThrGluCysGlyAsnGlyTYrVAlGluAlaGluGlu 20
DB 2502 CCAACAAAGCTATTGACGCCACGAGATGTGAAAGCATGCTGGAAGCTGGAGAG 2561
QY 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysCysLysCysSerLeuSer 40
DB 2562 TGTGATTGTGTTTTCATGTGATGCTATGATTTGCTGTAAAGAAATGTTCCCTCC 2621
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60
DB 2622 AACGGGGCTCAGTCAGCGACGAGCGGCGCTCTGTAAACAATACCTCATGTTTTCAGCCA 2681
QY 61 ATGGATYYTGluCysArgAspAlaValAsnGluCysAsp11ethrGluTYrCysThrGly 80
DB 2682 CAGGGTATGAATGCGGGATGCTGTGAACGAGTGTGATTTTCTGAATTTGTCTGGA 2741
QY 81 AspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrAlaCysAsnGlnAsn 100
DB 2742 GACTGTGTCAGTGCACCAATCTTCTAATAGCAAGACGATATGCATGCAATCAAAAT 2801
QY 101 GlnGlyArgCys 104
DB 2802 CAGGGCGGCTGC 2813

RESULT 7
US-10-260-506-9
; Sequence 9, Application US/10260506
; Publication No. US20030153064A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Talt, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/10/260,506
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410

```

```

; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 4720
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2129)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(4720)
; OTHER INFORMATION: n = A,T,C or G
US-10-260-506-9

Alignment Scores:
Pred. No.: 3,83e-69 Length: 4720
Score: 630.00 Matches: 104
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-634-252a-4_COPY_496_599 (1-104) x US-10-260-506-9 (1-4720)

QY 1 ProthrLysLeuPheGluProThrGluCysGlyAsnGlyTYrVAlGluAlaGluGlu 20
DB 1119 CCACAAAGCTATTGACGCCACGAGATGTGAAATGATGATCTGGAAGCTGGAGAGAG 1178
QY 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysCysLysCysSerLeuSer 40
DB 1179 TGTGATTGTGTTTTCATGTGATGCTATGATTTGCTGTAAAGAAATGTTCCCTCC 1238
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60
DB 1239 AACGGGGCTCAGTCAGCGACGAGCGGCGCTGTAAACAATACCTCATGTTTTCAGCCA 1298
QY 61 ATGGATYYTGluCysArgAspAlaValAsnGluCysAsp11ethrGluTYrCysThrGly 80
DB 1299 CAGGGTATGAATGCGGGATGCTGTGAACGAGTGTGATTTTCTGAATTTGTCTGGA 1358
QY 81 AspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTYrAlaCysAsnGlnAsn 100
DB 1359 GACTGTGTCAGTGCACCAATCTTCTAATAGCAAGACGATATGCATGCAATCAAAAT 1418
QY 101 GlnGlyArgCys 104
DB 1419 CAGGGCGGCTGC 1430

RESULT 8
US-09-918-995-22101
; Sequence 22101, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22101
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature

```

LOCATION: (1)...(494)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-22101

## Alignment Scores:

Pred. No.:	6,29e-69	Length:	494
Score:	618.00	Matches:	102
Percent Similarity:	98.08%	Conservative:	0
Best Local Similarity:	98.08%	Mismatches:	2
Query Match:	98.10%	Indels:	0
DB:	11	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-918-995-22101 (1-494)

Qy 1 ProThrLysLeuPheGluProThrGluCysGlyAsnGlyTyrrValGluAlaGlyGluGlu 20  
Db 126 CCAACAAGCTATTGAGCCCAACGGAATGTGAAATGATACGTTGAAAGCTGGGAGAG 185  
Qy 21 CysAspCysGlyPheHisValGluCysTyrrGlyLeuCysGlyLysCysSerLeuSer 40  
Db 186 TGTGATTGTGTTTTCATGATGGAATGCTATGATATCTGTAAATAATGTTCCCTCC 245  
Qy 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGluPro 60  
Db 246 AACGGGCTCACTGACGACGAGCGGCTGCTGTAAACAATACCTCAATGTTTTCAGCCA 305  
Qy 61 ArgGlyTyrrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrrCysThrGly 80  
Db 306 CGAGGGTATGAAATGCGGGATGCTGTGACGAGTCCGATATTACGAAATATTGACTGGA 365  
Qy 81 AspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTyrrAlaCysAsnGlnAsn 100  
Db 366 GACTCTGCTCAGTGCACCAACNAATCTTCATAGCAAGCGATAAGCATGCAATCAAAAT 425  
Qy 101 GlnGlyArgCys 104  
Db 426 CATGCCGCTGC 437

## RESULT 9

US-09-792-2008-17  
Sequence 17, Application US/097922008  
Patent No. US20020042368A1  
GENERAL INFORMATION:  
APPLICANT: Immunex Corporation  
APPLICANT: Fanslow, William C.  
APPLICANT: Poindexter, Kurt  
APPLICANT: Cerretti, Douglas P.  
TITLE OF INVENTION: INTEGRIN ANTAGONISTS  
FILE REFERENCE: 2958-A  
CURRENT APPLICATION NUMBER: US/09/792,2008  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 60/184,865  
PRIOR FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 1668  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fusion Construct  
NAME/KEY: CDS  
LOCATION: (25)...(1647)  
OTHER INFORMATION:  
US-09-792-2008-17

## Alignment Scores:

Pred. No.:	3,51e-64	Length:	1668
Score:	586.00	Matches:	96
Percent Similarity:	98.97%	Conservative:	0
Best Local Similarity:	98.97%	Mismatches:	1
Query Match:	93.02%	Indels:	0

DB: 9 Gaps: 0  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-792-2008-17 (1-1668)

Qy 8 ThrGluCysGlyAsnGlyTyrrValGluAlaGlyGluCysAspCysGlyPheHisVal 27  
Db 85 ACTAGTGTGAAATGATGATGCTGGAAGCTGGGAGAGTGTGATGTTTCATGTC 144  
Qy 28 GluCysTyrrGlyLeuCysCysLysLysCysSerLeuSerAsnGlyAlaHisCysSerAsp 47  
Db 145 GAATGCTATGATTTATCTCTTAAGAAATGTTTCCCTCCCAACGGGCTCACTGACGAGAC 204  
Qy 48 GlyProCysCysAsnAsnThrSerCysLeuPheGluProArgGlyTyrrGluCysArgAsp 67  
Db 205 GGGCCCTGCTGTAAACAATACCTCATGCTTTTCAGCAGAGGATGATGCGGGAT 264  
Qy 68 AlValAsnGluCysAspIleThrGluTyrrCysThrGlyAspSerGlyGlnCysProPro 87  
Db 265 GCTGTGAACGAGTGTGATTTACTGATATTGTAATGTAAGACCTGTGCTCAGTCCACCA 324  
Qy 88 AsnLeuHisLysGlnAspGlyTyrrAlaCysAsnGlnAsnGlnArgCys 104  
Db 325 AATCTTATTAAGCAAGCGATATGCTCAATCAAAATACAGGCCCTGC 375

## RESULT 10

US-09-809-790-3  
Sequence 3, Application US/09809790  
Patent No. US20020072102A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Balindur, Nand  
APPLICANT: Belsner, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/809,790  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 2088  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: zdin1 amino acid degenerate sequence  
NAME/KEY: variation  
LOCATION: (1)...(2088)  
OTHER INFORMATION: n is any nucleotide  
NAME/KEY: misc feature  
LOCATION: (1)...(2088)  
OTHER INFORMATION:  
US-09-809-790-3

## Alignment Scores:

Pred. No.:	1,19e-61	Length:	2088
Score:	567.00	Matches:	91
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	13
Query Match:	90.00%	Indels:	0
DB:	9	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-809-790-3 (1-2088)

Qy 1 ProThrLysLeuPheGluProThrGluCysGlyAsnGlyTyrrValGluAlaGlyGluGlu 20  
Db 1117 CCNACNAARYTNTTYGARCACNAGARTGYGNAAYCGNTAYGTNGARGCNGGARGAR 1176  
Qy 21 CysAspCysGlyPheHisValGluCysTyrrGlyLeuCysCysLysLysCysSerLeuSer 40  
Db 1177 TGGAGATGYGANTTYCAVGTNGARGTAYAGVNTGTGYTAAARAARTGYWSNYTWSN 1236

```

QY      41 AaNGlYAlaHicYsSeraspGlyProCysCyAaSnAnThrSerCysLeupheGlnPro 60
      1237 AAyGGNGcNCAYTGyWSNGAYGGNCcNTGTGYAAyAAyAACmWSNtGYTNTTYCARcCN 1296
QY      61 ArgGlyTYrGluCyAaArgaPaIaValaNGluCyAaSPlleThrGluTYrCysThrgly 80
      1297 MANGNTAYGARTGYMGNGAYGcNGTNAAYGARTGYGAyATHACNGARTAYTGyACNGcN 1356
QY      81 AspSerGlyGlnCySPProPcAaSnLeuHlslySGlnaSPGlyTYrAlaCysaNGlnaSn 100
      1357 GAyWSNGcNCARTGYcNCcNAAYTNCAYAAARcARGAyGNTAYGcNTGYAAyCARAAy 1416
Db      101 GlnGlyArgCys 104
      1417 CARGcNMNGTGY 1428

```

Search completed: October 21, 2003, 14:49:50  
 Job time : 271.906 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:27:07 ; Search time 3538.62 Seconds  
(without alignments)  
974.684 Million cell updates/sec

Title: US-09-634-252A-4\_COPY\_496\_599  
Perfect score: 630  
Sequence: 1 PTKLFEPTECGNGYVEAGE.....CPNLHKQDYACNQNQRC 104

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 10.0 , Fgapext 0.5  
Delop 6.0 , Delext 7.0

Searched: 3336368 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 6672376

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
O=/cgn2\_1/USPTO\_epool/US0964422/runat\_21102003\_090326\_3436/app\_query.fasta\_1.462  
-DB=Pending\_Patents\_NA\_Main -QFT=fastap -SUFFIX=p2n.inpm -MINMATCH=0.1  
-LOOPCL=0 -DOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US096334252.0CGN\_1\_1\_5064@runat\_21102003\_090326\_3436  
-NCPU=6 -ICPU=3 -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSELOCK=100  
-LONLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOPE=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELIXT=7

Database :

1: Pending Patents\_NA\_Main:\*  
2: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq.o.d:\*  
4: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US092A\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US092B\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US093A\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US093B\_COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*

29: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US099D\_COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US099E\_COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US099F\_COMB.seq:\*  
44: /cgn2\_6/ptodata/2/pna/US100A\_COMB.seq:\*  
45: /cgn2\_6/ptodata/2/pna/US100B\_COMB.seq:\*  
46: /cgn2\_6/ptodata/2/pna/US101A\_COMB.seq:\*  
47: /cgn2\_6/ptodata/2/pna/US101B\_COMB.seq:\*  
48: /cgn2\_6/ptodata/2/pna/US102A\_COMB.seq:\*  
49: /cgn2\_6/ptodata/2/pna/US102B\_COMB.seq:\*  
50: /cgn2\_6/ptodata/2/pna/US103A\_COMB.seq:\*  
51: /cgn2\_6/ptodata/2/pna/US103B\_COMB.seq:\*  
52: /cgn2\_6/ptodata/2/pna/US104A\_COMB.seq:\*  
53: /cgn2\_6/ptodata/2/pna/US104B\_COMB.seq:\*  
54: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*  
55: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*  
56: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*  
57: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*  
58: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*  
59: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*  
60: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*  
61: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*  
62: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*  
63: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*  
64: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*  
65: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*  
66: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*  
67: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*  
68: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*  
69: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*  
70: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*  
71: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*  
72: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*  
73: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*  
74: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*  
75: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*  
76: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*  
77: /cgn2\_6/ptodata/2/pna/US6023A\_COMB.seq:\*  
78: /cgn2\_6/ptodata/2/pna/US6023B\_COMB.seq:\*  
79: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*  
80: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*  
81: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*  
82: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*  
83: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*  
84: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*  
85: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*  
86: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*  
87: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*  
88: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq:\*  
89: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq:\*  
90: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq:\*  
91: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq:\*  
92: /cgn2\_6/ptodata/2/pna/US6037\_COMB.seq:\*  
93: /cgn2\_6/ptodata/2/pna/US6038\_COMB.seq:\*  
94: /cgn2\_6/ptodata/2/pna/US6039\_COMB.seq:\*  
95: /cgn2\_6/ptodata/2/pna/US6040\_COMB.seq:\*  
96: /cgn2\_6/ptodata/2/pna/US6041\_COMB.seq:\*  
97: /cgn2\_6/ptodata/2/pna/US6042\_COMB.seq:\*  
98: /cgn2\_6/ptodata/2/pna/US6043\_COMB.seq:\*  
99: /cgn2\_6/ptodata/2/pna/US6044\_COMB.seq:\*  
100: /cgn2\_6/ptodata/2/pna/US6045\_COMB.seq:\*  
101: /cgn2\_6/ptodata/2/pna/US6046\_COMB.seq:\*

102: /cgn2\_6/plodata/2/pna/US6047\_COMB.seq.\*  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	630	100.0	2268	27	US-09-631-534-1	Sequence 1, Appli
2	630	100.0	2268	34	US-09-809-617-1	Sequence 1, Appli
3	630	100.0	2268	34	US-09-809-790-1	Sequence 1, Appli
4	630	100.0	2268	49	US-10-260-506-1	Sequence 1, Appli
5	630	100.0	2394	40	US-09-949-016-5332	Sequence 5332, Ap
6	630	100.0	2489	27	US-09-634-252A-2	Sequence 2, Appli
7	630	100.0	2489	48	US-10-202-675-2	Sequence 2, Appli
8	630	100.0	2899	47	US-10-170-235-22436	Sequence 22436, A
9	630	100.0	2959	21	US-09-496-321-82	Sequence 82, Appli
10	630	100.0	2959	21	US-09-496-321-109	Sequence 109, App
11	630	100.0	2959	65	US-60-118-905-82	Sequence 82, Appli
12	630	100.0	2959	65	US-60-118-905-109	Sequence 109, App
13	630	100.0	3054	46	US-10-126-052A-149	Sequence 149, App
14	630	100.0	3054	46	US-10-126-052A-324	Sequence 324, App
15	630	100.0	4043	34	US-09-824-129-2	Sequence 2, Appli
16	630	100.0	4043	51	US-10-359-464-2	Sequence 2, Appli
17	630	100.0	4440	33	US-09-760-484-15	Sequence 15, Appli
18	630	100.0	4720	27	US-09-631-534-9	Sequence 9, Appli
19	630	100.0	4720	49	US-10-260-506-9	Sequence 9, Appli
20	630	100.0	6391	71	US-60-172-360-21719	Sequence 21719, A
21	630	100.0	6400	42	US-09-976-594-61	Sequence 61, Appli
22	630	100.0	6400	79	US-60-240-409-61	Sequence 61, Appli
23	630	100.0	6400	87	US-60-324-185-1573	Sequence 1573, Ap
24	618	98.1	494	17	US-09-234-611-16347	Sequence 16347, A
25	618	98.1	494	17	US-09-235-076-22101	Sequence 22101, A
26	618	98.1	494	18	US-09-289-768-20173	Sequence 20173, A
27	618	98.1	494	19	US-09-332-782-22101	Sequence 22101, A
28	618	98.1	494	32	US-09-737-223-22101	Sequence 22101, A
29	618	98.1	494	38	US-09-904-809-16347	Sequence 16347, A
30	618	98.1	494	38	US-09-918-995-22101	Sequence 22101, A
31	618	98.1	494	39	US-09-939-397-20173	Sequence 20173, A
32	617	97.9	6401	75	US-60-213-359-127	Sequence 127, App
33	611	97.0	2271	46	US-10-144-771-15737	Sequence 15737, A
34	611	97.0	2271	91	US-60-360-207-15737	Sequence 15737, A
35	586	93.0	1668	33	US-09-792-200B-17	Sequence 17, Appli
36	586	93.0	1668	33	US-09-792-200C-17	Sequence 17, Appli
37	567	90.0	2088	27	US-09-631-534-3	Sequence 3, Appli
38	567	90.0	2088	34	US-09-809-617-3	Sequence 3, Appli
39	567	90.0	2088	34	US-09-809-790-3	Sequence 3, Appli
40	567	90.0	2088	49	US-10-260-506-3	Sequence 3, Appli
41	567	90.0	2127	27	US-09-631-534-11	Sequence 11, Appli
42	567	90.0	2127	49	US-10-260-506-11	Sequence 11, Appli
43	551	87.5	468	22	US-09-519-705-803	Sequence 803, App
44	551	87.5	468	25	US-09-574-454-803	Sequence 803, App
45	551	87.5	468	48	US-10-221-279-792	Sequence 792, App

## ALIGNMENTS

RESULT 1  
 US-09-631-534-1  
 ; Sequence 1, Application US/09631534  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shepard, Paul O.  
 ; APPLICANT: Baidur, Nand  
 ; APPLICANT: Delsher, Theresa A.  
 ; APPLICANT: Bishop, Paul D.  
 ; APPLICANT: Taft, David W.  
 ; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP  
 ; FILE REFERENCE: 98-29C1  
 ; CURRENT APPLICATION NUMBER: US/09/631,534  
 ; CURRENT FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: 60/092,371  
 ;; PRIOR FILING DATE: 1998-07-10  
 ;; PRIOR APPLICATION NUMBER: 60/147, 410  
 ;; PRIOR FILING DATE: 1999-08-05  
 ;; PRIOR APPLICATION NUMBER: 09/351,414  
 ;; PRIOR FILING DATE: 1999-07-09  
 ;; NUMBER OF SEQ ID NOS: 15  
 ;; SOFTWARE: FastSeq for Windows Version 3.0  
 ;; SEQ ID NO 1  
 ;; LENGTH: 2268  
 ;; TYPE: DNA  
 ;; ORGANISM: Homo sapiens  
 ;; FEATURE:  
 ;; NAME/KEY: CDS  
 ;; LOCATION: (3)...(2090)  
 ;; NAME/KEY: misc\_feature  
 ;; LOCATION: (1)...(2268)  
 ;; OTHER INFORMATION: n = A,T,C or G  
 US-09-631-534-1

## Alignment Scores:

Pred. No.: 4,426-52 Length: 2268  
 Score: 630.00 Matches: 104  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 27 Gaps: 0

US-09-634-252A-4\_COPY\_496\_599 (1-104) x US-09-631-534-1 (1-2268)

QY 1 Prothrlvsleuphegluprothrlgucysglvaenglytyrvalglualglglu 20  
 DB 1119 CCACAAAGCATTTGAGCCACCGAATGTGAATGATACGTGAGCTGGAGAG 1178  
 QY 21 CysaarpCysglvPhhisvalglucystyrglyleucyCysglvlyCysSerleuser 40  
 DB 1179 TGTGATTGTGTTTCATGTGAATGATGATGATTATCTGAAGAAATGTTCCCTTC 1238  
 QY 41 AenglylAHIScysSerAspGlyProCysCysAsnaanthrSerCysLeupheglupro 60  
 DB 1239 AACGGGCTCACTGACGAGCGGCGCTGCTGTAACTACTGCTTTTAAAGCA 1298  
 QY 61 ArgglyTYrgucysargaspalaValaenglyCysaepiliethtgutyrcysThrGly 80  
 DB 1299 CGAGGTATGATGATCGCGGATGCTGTGAACGATGTGATATTAAGATATTGACTGA 1358  
 QY 81 AppserGlyGlnCysProProAsnleuhtslysglnaepglyTYrAlaCysasnlnasn 100  
 DB 1359 GACTGTGTCAGTGTCCACCAATCTTCATAAGCAAGCATATGCAATCAAAAT 1418  
 QY 101 GlnGlyArGcys 104  
 DB 1419 CAGGGCCGCTGC 1430

RESULT 2  
 US-09-809-617-1  
 ; Sequence 1, Application US/09809617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shepard, Paul O.  
 ; APPLICANT: Baidur, Nand  
 ; APPLICANT: Delsher, Theresa A.  
 ; APPLICANT: Bishop, Paul D.  
 ; TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
 ; FILE REFERENCE: 98-29  
 ; CURRENT APPLICATION NUMBER: US/09/809,617  
 ; CURRENT FILING DATE: 2001-03-15  
 ; PRIOR APPLICATION NUMBER: US/09/351,414  
 ; PRIOR FILING DATE: 1999-07-09  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2268  
 ; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-09-809-617-1

## Alignment Scores:

Pred. No.:	4,42e-52	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	34	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-809-617-1 (1-2268)

QY 1 ProThrIysLeuPheGluProThrGluCysGlyAsnGlyTyryValGluAlaGlyGluGlu 20  
Db 1119 CCAACAAGCTATTGAGCCACGGAATGTGAATGATACCTGGAAGCTGGGAGAG 1178  
QY 21 CysAspCysGlyPheHisValGluCysTyryGlyLeuCysCysIysIysCysSerLeuSer 40  
Db 1179 TGTGATTGTGTTTTCATGTGGAATGCTATGATTATGCTGTAGAAATGTTCCCTCC 1238  
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAnthrSerCysLeuPheGluPro 60  
Db 1239 AACGGGCTCCTCAGCGAGGAGGCGGCTGCTGTAACAAATCCTATGCTTTTCACCA 1298  
QY 61 ArgGlyTyryGluCysArgAspAlaValAsnGluCysAspIleThrGluTyryCysThrGly 80  
Db 1299 CGAGGATGATGAATCCCGGATGCTGTGAACGAGTGTATTAATTAATTTGACTGGA 1358  
QY 81 AspSerGlyGlnCysProProAsnLeuHisIysGlnAspGlyTyryAlaCysAsnGlnAsn 100  
Db 1359 GACTCTGTGTCAGTCCACCAAAATCTTCATAGCAAGCGATGATGCAATCAAAAT 1418  
QY 101 GlnGlyArgCys 104  
Db 1419 CAGGGCCGCTGC 1430

## RESULT 3

US-09-809-790-1  
Sequence 1, Application US/09809790  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baindur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/809,790  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-09-809-790-1

## Alignment Scores:

Pred. No.:	4,42e-52	Length:	2268
------------	----------	---------	------

Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	34	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-809-790-1 (1-2268)

QY 1 ProThrIysLeuPheGluProThrGluCysGlyAsnGlyTyryValGluAlaGlyGluGlu 20  
Db 1119 CCAACAAGCTATTGAGCCACGGAATGTGAATGATACCTGGAAGCTGGGAGAG 1178  
QY 21 CysAspCysGlyPheHisValGluCysTyryGlyLeuCysCysIysIysCysSerLeuSer 40  
Db 1179 TGTGATTGTGTTTTCATGTGGAATGCTATGATTATGCTGTAGAAATGTTCCCTCC 1238  
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAnthrSerCysLeuPheGluPro 60  
Db 1239 AACGGGCTCCTCAGCGAGGAGGCGGCTGCTGTAACAAATCCTATGCTTTTCACCA 1298  
QY 61 ArgGlyTyryGluCysArgAspAlaValAsnGluCysAspIleThrGluTyryCysThrGly 80  
Db 1299 CGAGGATGATGAATCCCGGATGCTGTGAACGAGTGTATTAATTAATTTGACTGGA 1358  
QY 81 AspSerGlyGlnCysProProAsnLeuHisIysGlnAspGlyTyryAlaCysAsnGlnAsn 100  
Db 1359 GACTCTGTGTCAGTCCACCAAAATCTTCATAGCAAGCGATGATGCAATCAAAAT 1418  
QY 101 GlnGlyArgCys 104  
Db 1419 CAGGGCCGCTGC 1430

## RESULT 4

US-10-260-506-1  
Sequence 1, Application US/10260506  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baindur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Taft, David W.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP  
FILE REFERENCE: 98-29C1  
CURRENT APPLICATION NUMBER: US/10/260,506  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: 60/092,371  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/147,410  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-10-260-506-1

## Alignment Scores:

Pred. No.:	4,42e-52	Length:	2268
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	49	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-10-260-506-1 (1-2268)

QY 1 ProthrlvsleuphegluProthrgluCysglsyAsnglyTYrValglnlaaglyglu 20  
 DB 1119 CCACAAAGCTATTGAGCCACGGAATGTGAAATGATGAGAACTGGGAGAG 1178  
 QY 21 CysaspCysglsyPhehisValglnCysTYrTglsyLeuCyAslsyLysCysSerleuser 40  
 DB 1179 TGTGATGTGCTTTCATGTGAAATGCTATGATTTGCTGTAAAGAAATGTTCCCTCTCC 1238  
 QY 41 AenglyAlahisCysSerleuserProCyCysAsnaenthSerCysleuphegluPro 60  
 DB 1239 AACGGGGCTCAGTGCAGCGACGAGGCGCCCTCTGTAAACAATCCTCATGCTTTTCAGCCA 1298  
 QY 61 ArgglyTYrTglsyCysArgspAlaValaAsnglyCysAsp1lethglnlyrCysThrgly 80  
 DB 1299 CAGGGGTATGAATGCCGGATGCTGTGAACGAGTGTGATTAATGATGATGCTGGA 1358  
 QY 81 AepserglynCysProProAsnleuHisLysGlnAspGlyTYrAlaCysAsnglnasn 100  
 DB 1359 GACTCTGTCAGTGCAGTGCACCAATCTTCTAAGCAAGAGATATGCATCAATCAAAAT 1418  
 QY 101 GlnGlyArgCys 104  
 DB 1419 CAGGGCCGCTGC 1430

RESULT 5  
 US-09-949-016-5332

; Sequence 5332, Application US/09949016  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5332  
 ; LENGTH: 2394  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-5332

Alignment Scores:

Pred. No.:	4,716-52	Length:	2394
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	40	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-949-016-5332 (1-2394)

QY 1 ProthrlvsleuphegluProthrgluCysglsyAsnglyTYrValglnlaaglyglu 20  
 DB 1050 CCACAAAGCTATTGAGCCACGGAATGTGAAATGATGAGAACTGGGAGAG 1109  
 QY 21 CysaspCysglsyPhehisValglnCysTYrTglsyLeuCyAslsyLysCysSerleuser 40  
 DB 1110 TGTGATGTGCTTTCATGTGAAATGCTATGATTTGCTGTAAAGAAATGTTCCCTCTCC 1169  
 QY 41 AenglyAlahisCysSerleuserProCyCysAsnaenthSerCysleuphegluPro 60  
 DB 1170 AACGGGGCTCAGTGCAGCGACGAGGCGCCCTCTGTAAACAATCCTCATGCTTTTCAGCCA 1229

QY 61 ArgglyTYrTglsyCysArgspAlaValaAsnglyCysAsp1lethglnlyrCysThrgly 80  
 DB 1230 CAGGGGTATGAATGCCGGATGCTGTGAACGAGTGTGATTAATGATGATGCTGGA 1289  
 QY 81 AepserglynCysProProAsnleuHisLysGlnAspGlyTYrAlaCysAsnglnasn 100  
 DB 1290 GACTCTGTCAGTGCAGTGCACCAATCTTCTAAGCAAGAGATATGCATCAATCAAAAT 1349  
 QY 101 GlnGlyArgCys 104  
 DB 1350 CAGGGCCGCTGC 1361

RESULT 6  
 US-09-634-252a-2

; Sequence 2, Application US/09634252A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Coriell, Douglas P.  
 ; TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES  
 ; FILE REFERENCE: 03260.0051-00304  
 ; CURRENT APPLICATION NUMBER: US/09/634,252A  
 ; CURRENT FILING DATE: 2000-08-07  
 ; PRIOR APPLICATION NUMBER: 60/074,310  
 ; PRIOR FILING DATE: 1998-02-11  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 2499  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-634-252a-2

Alignment Scores:

Pred. No.:	4,956-52	Length:	2499
Score:	630.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	27	Gaps:	0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-634-252a-2 (1-2499)

QY 1 ProthrlvsleuphegluProthrgluCysglsyAsnglyTYrValglnlaaglyglu 20  
 DB 1486 CCACAAAGCTATTGAGCCACGGAATGTGAAATGATGAGAACTGGGAGAG 1545  
 QY 21 CysaspCysglsyPhehisValglnCysTYrTglsyLeuCyAslsyLysCysSerleuser 40  
 DB 1546 TGTGATGTGCTTTCATGTGAAATGCTATGATTTGCTGTAAAGAAATGTTCCCTCTCC 1605  
 QY 41 AenglyAlahisCysSerleuserProCyCysAsnaenthSerCysleuphegluPro 60  
 DB 1606 AACGGGGCTCAGTGCAGCGACGAGGCGCCCTCTGTAAACAATCCTCATGCTTTTCAGCCA 1665  
 QY 61 ArgglyTYrTglsyCysArgspAlaValaAsnglyCysAsp1lethglnlyrCysThrgly 80  
 DB 1666 CAGGGGTATGAATGCCGGATGCTGTGAACGAGTGTGATTAATGATGATGCTGGA 1725  
 QY 81 AepserglynCysProProAsnleuHisLysGlnAspGlyTYrAlaCysAsnglnasn 100  
 DB 1726 GACTCTGTCAGTGCAGCGACGAGGCGCCCTCTGTAAACAATCCTCATGATGATGCAATCAAAAT 1785  
 QY 101 GlnGlyArgCys 104  
 DB 1786 CAGGGCCGCTGC 1797

RESULT 7  
 US-10-202-675-2

; Sequence 2, Application US/10202675  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Coriell, Douglas P.  
 ; TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES  
 ; FILE REFERENCE: 03260.0051-00304  
 ; CURRENT APPLICATION NUMBER: US/10/202,675



;; CURRENT FILING DATE: 2002-07-23  
;; PRIOR APPLICATION NUMBER: US/09/634,252  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: 60/074,310  
;; PRIOR FILING DATE: 1998-02-11  
;; NUMBER OF SEQ ID NOS: 10  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 2  
;; LENGTH: 2499  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-202-675-2

Alignment Scores:  
Pred. No.: 4,95e-52 Length: 2499  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 48 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-10-202-675-2 (1-2499)

QY 1 Prothrluylsleuphegluprothrglucysgilyasncllytyryvalglualaglyglu 20  
Db 1486 CCAACAAAGCTATTGTGAGCCACGGAATGTGAATGATACGTGAAGCTGGGAGAG 1545

QY 21 CysaspCysgilyPhehlsvalglucyetyrglyleucysgilylucysseleuser 40  
Db 1546 TGTGATTGTGTTTTCATGTGGAATGCTATGATTTGCTTAAGAAATGTTCCCTCC 1605

QY 41 AsncllylalihsCysseAspGlyProCysCysasnAnthrserCysleupheglupro 60  
Db 1606 AACGGGGCTCACTGACGAGGAGGCGCTGTAACATACCTGATCTTTTCAGCA 1665

QY 61 ArgcllytyrglucysatgaspAlavalasnclucysapllethnglytyrcythrgly 80  
Db 1666 CGAGGATATGATATCCGGGATGCTGTGAACAGATGTGATTAATGATATTGACTGGA 1725

QY 81 AspserglylncysproProasnleuhslysglnaspglytyryalaCysasnGlnasn 100  
Db 1726 GACTCTGCTAGTCCGACCACTTCTATTAAGCAAGCGATATGATGCAATCAAAAT 1785

QY 101 GlncllyargCys 104  
Db 1786 CAGGGCCGCTGC 1797

RESULT 8  
US-10-170-235-22436  
;; Sequence 22436, Application US/10170235  
;; GENERAL INFORMATION:  
;; APPLICANT: VENTER, J. Craig  
;; TITLE OF INVENTION: KITS SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN  
;; FILE REFERENCE: CLO01380  
;; CURRENT APPLICATION NUMBER: US/10/170,235  
;; CURRENT FILING DATE: 2003-03-17  
;; NUMBER OF SEQ ID NOS: 42514  
;; SEQ ID NO 22436  
;; LENGTH: 2899  
;; TYPE: DNA  
;; ORGANISM: HUMAN  
US-10-170-235-22436

Alignment Scores:  
Pred. No.: 5.88e-52 Length: 2899  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 47 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-10-170-235-22436 (1-2899)

QY 1 Prothrluylsleuphegluprothrglucysgilyasncllytyryvalglualaglyglu 20  
Db 1555 CCAACAAAGCTATTGTGAGCCACGGAATGTGAATGATACGTGAAGCTGGGAGAG 1614

QY 21 CysaspCysgilyPhehlsvalglucyetyrglyleucysgilylucysseleuser 40  
Db 1615 TGTGATTGTGTTTTCATGTGGAATGCTATGATTTGCTTAAGAAATGTTCCCTCC 1674

QY 41 AsncllylalihsCysseAspGlyProCysCysasnAnthrserCysleupheglupro 60  
Db 1675 AACGGGGCTCACTGACGAGGAGGCGCTGTAACATACCTGATCTTTTCAGCA 1734

QY 61 ArgcllytyrglucysatgaspAlavalasnclucysapllethnglytyrcythrgly 80  
Db 1735 CGAGGATATGATATCCGGGATGCTGTGAACAGATGTATTTACTGAAATTTGACTGGA 1794

QY 81 AspserglylncysproProasnleuhslysglnaspglytyryalaCysasnGlnasn 100  
Db 1795 GACTCTGCTAGTCCGACCACTTCTATTAAGCAAGCGATATGATGCAATCAAAAT 1854

QY 101 GlncllyargCys 104  
Db 1855 CAGGGCCGCTGC 1866

RESULT 9  
US-09-496-321-82  
;; Sequence 82, Application US/09496321  
;; GENERAL INFORMATION:  
;; APPLICANT: Reddy, Roopa  
;; APPLICANT: Guegler, Karl, J.  
;; APPLICANT: Au-Young, Janice  
;; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING RECEPTOR PROTEINS  
;; FILE REFERENCE: PA-0014 US  
;; CURRENT APPLICATION NUMBER: US/09/496,321  
;; CURRENT FILING DATE: 2000-02-01  
;; EARLIER APPLICATION NUMBER: 60/118,905  
;; EARLIER FILING DATE: February 5, 1999  
;; NUMBER OF SEQ ID NOS: 125  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 82  
;; LENGTH: 2959  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: 305461  
US-09-496-321-82

Alignment Scores:  
Pred. No.: 6.03e-52 Length: 2959  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-496-321-82 (1-2959)

QY 1 Prothrluylsleuphegluprothrglucysgilyasncllytyryvalglualaglyglu 20  
Db 426 CCAACAAAGCTATTGTGAGCCACGGAATGTGAATGATGATGGAAGCTGGGAGAG 485.

QY 21 CysaspCysgilyPhehlsvalglucyetyrglyleucysgilylucysseleuser 40  
Db 486 TGTGATTGTGTTTTCATGTGGAATGCTATGATTTGCTTAAGAAATGTTCCCTCC 545

QY 41 AsncllylalihsCysseAspGlyProCysCysasnAnthrserCysleupheglupro 60  
Db 546 AACGGGGCTCACTGACGAGGAGGCGCTGTAACATACCTGATCTTTTCAGCA 605

QY 61 ArgcllytyrglucysatgaspAlavalasnclucysapllethnglytyrcythrgly 80  
Db 606 CGAGGATATGATATCCGGGATGCTGTGAACAGATGTGATTAATGATGACTGGA 665

QY 81 AspSerGlyGlnCysProProAsnLeuHisIysGlnAspGlyTyrAlaCysAsnGlnAsn 100  
 DB 666 GACTCTGGTCAGTGGCCCAACAAATCTTCAATAGCAAGACGATGATGCAATCAATAAT 725  
 QY 101 GINGLYARGCys 104  
 DB 726 CAGGGCCGCTGC 737

## RESULT 10

US-09-496-321-109  
 ; Sequence 109, Application US/09496321  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reddy, Roopa  
 ; APPLICANT: Guegler, Karl, J.  
 ; APPLICANT: Au-Young, Janice  
 ; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING RECEPTOR PROTEINS  
 ; FILE REFERENCE: PA-0014 US  
 ; CURRENT APPLICATION NUMBER: US/09/496,321  
 ; CURRENT FILING DATE: 2000-02-01  
 ; EARLIER APPLICATION NUMBER: 60/118,905  
 ; EARLIER FILING DATE: February 5, 1999  
 ; NUMBER OF SEQ ID NOS: 125  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 109  
 ; LENGTH: 2959  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE: -  
 ; OTHER INFORMATION: 3859516  
 US-09-496-321-109

## Alignment Scores:

Pred. No.: 6, 03e-52 Length: 2959  
 Score: 630.00 Matches: 104  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-496-321-109 (1-2959)

QY 1 ProThrIysLeuPheGluProThrGluCysGlyAsnGlyTyrValGlnAlaGlyGlu 20  
 DB 426 CCAACAAAGCTATTTCAGCCCAACGAATGTGAAATGATACGTGAACTGGGAGAG 485  
 QY 21 CysAspCysGlyPheHisValGluCysTyrGlyLeuCysCysIysIysCysSerLeuSer 40  
 DB 486 TGTGATTGTGGTTTTCATGTGAAATGCTAAGGATTAGCTGTAGAAATGTTCCCTCTCC 545  
 QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGlnPro 60  
 DB 546 AACGGGGCTCAGTCGACGACGCGGCCCTCTGTAAACATACCTCATGCTTTTTCAGCCA 605  
 QY 61 ArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGluTyrCysThrGly 80  
 DB 606 CGAGGGGTATGAATGCGGGATGCTGTGAACGAGTGTATTACTGAATATTGTACTGGA 665  
 QY 81 AspSerGlyGlnCysProProAsnLeuHisIysGlnAspGlyTyrAlaCysAsnGlnAsn 100  
 DB 666 GACTCTGGTCAGTGGCCCAACAAATCTTCAATAGCAAGACGATGATGCAATCAATAAT 725  
 QY 101 GINGLYARGCys 104  
 DB 726 CAGGGCCGCTGC 737

Search completed: October 21, 2003, 14:37:17  
 Job time : 3546.62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:08:40 ; Search time 17.6604 Seconds  
(without alignments)  
190.928 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599

Perfect score: 630  
Sequence: 1 PTKLFEPTECGNGYVGEAGE.....CPNHLKODGVACNONGRC 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 153893 seqs, 32421815 residues

Total number of hits satisfying chosen parameters: 153893

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCIT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	832	7	US-60-495-114-1465 Sequence 1465, Ap
2	630	100.0	832	7	US-60-495-114-1466 Sequence 1466, Ap
3	404.5	64.2	369	7	US-60-495-114-1961 Sequence 1961, Ap
4	404.5	64.2	623	7	US-60-487-610-2361 Sequence 2361, Ap
5	404.5	64.2	623	7	US-60-485-450-1482 Sequence 1482, Ap
6	404.5	64.2	623	7	US-60-495-114-1962 Sequence 1962, Ap
7	404.5	64.2	769	7	US-60-487-610-2360 Sequence 2360, Ap
8	404.5	64.2	769	7	US-60-485-450-1481 Sequence 1481, Ap
9	404.5	64.2	769	7	US-60-495-114-1963 Sequence 1963, Ap
10	404.5	64.2	773	7	US-60-495-114-1964 Sequence 1964, Ap
11	404.5	64.2	775	7	US-60-487-610-2362 Sequence 2362, Ap
12	404.5	64.2	775	7	US-60-485-450-1483 Sequence 1483, Ap
13	363.5	57.7	761	7	US-60-495-114-1301 Sequence 1301, Ap
14	363.5	57.7	761	7	US-60-495-114-1302 Sequence 1302, Ap
15	363.5	57.7	823	7	US-60-490-890-1043 Sequence 1043, Ap
16	363.5	57.7	823	7	US-60-495-114-1304 Sequence 1304, Ap
17	363.5	57.7	823	7	US-60-495-114-1309 Sequence 1309, Ap
18	363.5	57.7	853	7	US-60-495-114-1310 Sequence 1310, Ap
19	363.5	57.7	859	7	US-60-490-890-1039 Sequence 1039, Ap
20	363.5	57.7	864	7	US-60-495-114-1308 Sequence 1308, Ap
21	363.5	57.7	870	7	US-60-490-890-1041 Sequence 1041, Ap
22	363.5	57.7	870	7	US-60-495-114-1305 Sequence 1305, Ap
23	363.5	57.7	870	7	US-60-495-114-1311 Sequence 1311, Ap
24	363.5	57.7	893	7	US-60-495-114-1307 Sequence 1307, Ap
25	363.5	57.7	899	7	US-60-495-114-1303 Sequence 1303, Ap
26	363.5	57.7	899	7	US-60-495-114-1306 Sequence 1306, Ap

27	363.5	57.7	900	7	US-60-495-114-1312 Sequence 1312, Ap
28	292	46.3	746	6	US-10-670-184-4 Sequence 4, Appl
29	292	46.3	787	6	PCT-US03-30720-951 Sequence 951, Appl
30	292	46.3	787	6	US-10-670-184-5 Sequence 5, Appl
31	292	46.3	787	7	US-60-495-114-2100 Sequence 2100, Ap
32	292	46.3	813	7	US-60-495-114-2101 Sequence 2101, Ap
33	292	46.3	824	1	PCT-US03-28227-4610 Sequence 4610, Ap
34	292	46.3	824	1	PCT-US03-28227-4612 Sequence 4612, Ap
35	292	46.3	840	1	PCT-US03-30720-1275 Sequence 1275, Ap
36	292	46.3	849	6	US-10-670-184-6 Sequence 6, Appl
37	292	46.3	863	1	PCT-US03-28227-4611 Sequence 4611, Ap
38	274	43.5	356	7	US-60-495-114-1584 Sequence 1584, Ap
39	274	43.5	358	1	PCT-US03-26780-1448 Sequence 1448, Ap
40	274	43.5	409	7	US-60-495-114-1583 Sequence 1583, Ap
41	274	43.5	652	7	US-60-495-114-1585 Sequence 1585, Ap
42	274	43.5	778	7	US-60-487-610-1926 Sequence 1926, Ap
43	274	43.5	781	7	US-60-487-610-1924 Sequence 1924, Ap
44	274	43.5	802	7	US-60-487-610-1925 Sequence 1925, Ap
45	274	43.5	814	1	PCT-US03-28751-4 Sequence 4, Appl

## ALIGNMENTS

```
RESULT 1
US-60-495-114-1465
; Sequence 1465, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1465
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-495-114-1465

Query Match      100.0%; Score 630; DB 7; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PTKLFEPTECGNGYVGEAGECDGFFHVECYGLCKCKSLNSGAHCSDPCCNNTSCUFOP 60
      |||
Db      496 PTKLFEPTECGNGYVGEAGECDGFFHVECYGLCKCKSLNSGAHCSDPCCNNTSCUFOP 555

QY      61 RGVECRDANVECDITTEYCTGDSGCCPPNHLKODGVACNONGRC 104
      |||
Db      556 RGVECRDANVECDITTEYCTGDSGCCPPNHLKODGVACNONGRC 599

RESULT 2
US-60-495-114-1466
; Sequence 1466, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1466
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
```



```

RESULT 7
US-60-487-610-2360
: Sequence 2360, Application US/60487610
:
: GENERAL INFORMATION:
:
: APPLICANT: CARGILL, Michele
:
: APPLICANT: HUANG, Hongjin
:
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
:
: TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
:
: TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
:
: FILE REFERENCE: C1001469
:
: CURRENT APPLICATION NUMBER: US/60/487,610
:
: CURRENT FILING DATE: 2003-07-17
:
: NUMBER OF SEQ ID NOS: 97101
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 2360
:
: LENGTH: 769
:
: TYPE: PR1
:
: ORGANISM: Homo sapiens
:
: US-60-487-610-2360

```

Query Match	64.2%	Score 404.5:	DB 7,	Length 769;
Best Local Similarity	65.4%;	Pred. No. 3.8e-11;		
Matches	70;	Conservative	9; Mismatches	23; Indels 5; Gaps 2
Oy	1	PTKLPEPFGCGNGYEAGEECDCGFHVECY---	GLCKKCSLSNIAHSDEPCNNISCL	57
	:			
Db	438	PLKLLDPEPCGNGFYEAEGEEDCCGGVQESRAGNCCKKCTLTTHAMCSDGLCCR--CK	495	
Oy	58	FQPRGYECRDANNECDITEYCTGDSGCCPPNLHKDDGVACNQGRIC	104	
	::			
Db	496	YEPKVCSCREAVNECDIAETCTGDSGCCPPNLHKDDGVACDHEQRIC	542	

```

RESULT 8
US-60-485-450-1481
: Sequence 1481, Application US/60485450
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: CHANG, Sheng-Yung
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
: TITLE OF INVENTION: VIRUS-INJECTED SUBJECTS, METHODS OF DETECTION AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CL001470
: CURRENT APPLICATION NUMBER: US/60/485,450
: CURRENT FILING DATE: 2003-07-09
: NUMBER OF SEQ. ID NOS: 47859
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1481
: LENGTH: 769
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-485-450-1481

```

Query Match	54.2%	Score 404.5;	DB 7;	length 769;
Best Local Similarity	65.4%	Pred. No. 3.8e-11;		
Matches	70;	Conservative	9;	Mismatches 22;
			Indels	5;
			Gaps	2;
Qy	1	PTKLEPPEPCNGYVEAGEECDCGFHYCY---	GLCKCKKSLISNCAHSDPEPCNTTCL	57
Dy	438	PLKLLDPEPCNGFVEAESEECDCGSSVQEC	SRAGNCKCKCILLHMLMSDGLCLR--CK	495
Qy	58	FQPRGYECRDVAVNECDITETCTGDSGCC	CPPLHKKODGYACANONCGR	104
Dy	496	YEPKRVSCREAVNECDIAETCTGDSGCC	CPPLHKKIDGYCYCHDEQRC	342

RESULT 9  
US-60-495-114-1963  
Sequence 1963, Application US/60495114  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

```

: TITLE OF INVENTION:  ENCODING HUMAN PROTEASE PROTEINS,  METHODS OF DETECTION AND
:
: TITLE OF INVENTION:  USES THEREOF
: FILE REFERENCE:  CL001480
: CURRENT APPLICATION NUMBER:  US/60/495,114
: CURRENT FILING DATE:  2003-08-15
: NUMBER OF SEQ ID NOS:  91238
: SOFTWARE:  FastSeq for Windows Version 4.0
: SEQ ID NO 1963
: LENGTH:  769
: TYPE:  prt
: ORGANISM:  Homo sapiens
: US-60-495-114-1963

```

Query Match	64.2%	Score	404.5	DB	7	Length	769
Best Local Similarity	65.4%	Pred.	No. 3.8e-11				
Matches	70	Conservative	9	Mismatches	23	Indels	5
				Gaps	2		

  

OY	1	PTKLFEPTGCGNGVYAGEECDCGFHNECY----	GLCKCKSLGSAHCSDPGCCNNTSCL	57
Db	438	PLKLIDPECGNGVYAGEECDCSSVBCSAGSNCKKCTLT	THDAMCSDLCCRR--CK	495
OY	58	FQPRGVCRDVAVNECDITTEYCTGDSGCCPPLHLKODG	YACANQNGRC	104
Db	496	YEPFGVSCREAVNECDIAETGTGSSQCPPLHLKIDG	YCTCHDEGRC	542

```

RESULT 10
US-60-495-114-1964
; Sequence 1964, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CU001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1964
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-495-114-1964

```

Query Match	64.2%	Score 404.5	DB 7	Length 773
Best Local Similarity	65.4%	Pred. No. 3.8e-11		
Matches	70	Conservative 9	Mismatches 23	Indels 5
Gaps	2			
Oy	1	PTKLPEPECCNGVGEAGECCDGGPFHVCY----	GLCKCKGSLNSGAHCSDPGCCNNTSCL	57
Db	438	PLKLDPCECCNGVEAGEBECDCGSVOCSSRAGNCCCKCTLTTHDAMCSDDGLCCRR--CK		495
Oy	58	FOPRGVECDPAVNECDITTEYCTGDSGOCPPYLTHKODGYACNQNQRC		104
Db	496	YEPGVGSCREAVNECDIAETCTGSSQCCPPYLHLKLDGYICDHQRC		542

Search completed: October 21, 2003, 10:21:22  
Job time : 18.6604 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:06:10 ; Search time 314.616 Seconds  
(without alignments)  
300.784 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599

Perfect score: 630  
Sequence: 1 PTLFPEPTCEGNGVEAGE.....CPNLHKDGVACHONGRC 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/US06\_PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097A\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US097B\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/2/paa/US099A\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/2/paa/US102\_COMB.pep.\*
- 29: /cgn2\_6/ptodata/2/paa/US103\_COMB.pep.\*
- 30: /cgn2\_6/ptodata/2/paa/US104\_COMB.pep.\*
- 31: /cgn2\_6/ptodata/2/paa/US106\_COMB.pep.\*
- 32: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	686	24	US-09-949-016-11203
					Sequence 11203, A

2	630	100.0	636	20	US-09-631-534-2	Sequence 2, App1
3	630	100.0	636	23	US-09-809-617-2	Sequence 2, App1
4	630	100.0	696	23	US-09-809-790-2	Sequence 2, App1
5	630	100.0	636	28	US-10-260-506-2	Sequence 2, App1
6	630	100.0	709	20	US-09-631-534-10	Sequence 10, App1
7	630	100.0	709	28	US-10-260-506-10	Sequence 10, App1
8	630	100.0	776	22	US-09-760-484-315	Sequence 315, App
9	630	100.0	832	22	US-09-634-252a-4	Sequence 4, App1
10	630	100.0	832	22	US-09-791-537-64795	Sequence 64795, A
11	630	100.0	832	22	US-09-824-129-3	Sequence 3, App1
12	630	100.0	832	27	US-10-126-052A-150	Sequence 150, App
13	630	100.0	832	27	US-10-126-052A-325	Sequence 325, App
14	630	100.0	832	29	US-10-202-675-4	Sequence 4, App1
15	630	100.0	832	29	US-10-359-464-3	Sequence 3, App1
16	611	97.0	829	22	US-09-791-537-142799	Sequence 142799, A
17	586	93.0	540	22	US-09-792-200B-18	Sequence 18, App1
18	586	93.0	540	22	US-09-792-200C-18	Sequence 18, App1
19	553	87.8	154	28	US-10-221-279-6972	Sequence 6972, App
20	515.5	81.8	368	22	US-09-760-484-477	Sequence 477, App
21	515.5	81.8	368	28	US-10-211-364-1048	Sequence 1048, App
22	515.5	81.8	368	28	US-10-216-893-296	Sequence 296, App
23	515.5	81.8	368	28	US-10-217-651-431	Sequence 431, App
24	404.5	64.2	518	22	US-09-791-537-19102	Sequence 19102, A
25	404.5	64.2	524	22	US-09-791-537-88445	Sequence 88445, A
26	404.5	64.2	623	32	US-60-453-050-13203	Sequence 13203, A
27	404.5	64.2	623	32	US-60-453-135-13203	Sequence 13203, A
28	404.5	64.2	623	32	US-60-466-412-13203	Sequence 13203, A
29	404.5	64.2	664	22	US-09-791-537-107561	Sequence 107561, A
30	404.5	64.2	769	22	US-09-791-537-137085	Sequence 137085, A
31	404.5	64.2	769	24	US-09-949-016-9605	Sequence 9605, App
32	404.5	64.2	769	24	US-09-949-016-9606	Sequence 9606, App
33	404.5	64.2	769	32	US-60-453-050-13202	Sequence 13202, A
34	404.5	64.2	769	32	US-60-453-135-13202	Sequence 13202, A
35	404.5	64.2	769	32	US-60-466-412-13202	Sequence 13202, A
36	404.5	64.2	773	22	US-09-791-537-137203	Sequence 137203, A
37	404.5	64.2	775	32	US-60-453-050-13204	Sequence 13204, A
38	404.5	64.2	775	32	US-60-453-135-13204	Sequence 13204, A
39	404.5	64.2	775	32	US-60-466-412-13204	Sequence 13204, A
40	404.5	64.2	785	32	US-60-212-656-502	Sequence 502, App
41	404.5	64.2	785	32	US-60-230-435-1310	Sequence 1310, App
42	404.5	64.2	808	32	US-60-207-315-350	Sequence 352, App
43	404.5	64.2	858	32	US-60-212-656-480	Sequence 480, App
44	404.5	64.2	858	32	US-60-230-435-1465	Sequence 1465, App
45	376.5	59.8	664	32	US-60-230-435-1070	Sequence 1070, App

#### ALIGNMENTS

RESULT 1  
US-09-949-016-11203  
Sequence 11203, Application US/09949016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: C0001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 11203  
LENGTH: 686  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-11203

Query Match 100.0%; Score 630; DB 24; Length 686;  
Best Local Similarity 100.0%; Pred. No. 6.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60

DB 350 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 409

QY 61 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 104

DB 410 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 453

## RESULT 2

US-09-631-534-2  
Sequence 2, Application US/09631534

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Baidur, Nand

APPLICANT: Deisher, Theresa A.

APPLICANT: Bishop, Paul D.

APPLICANT: Taft, David W.

TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP

FILE REFERENCE: 98-29C1

CURRENT APPLICATION NUMBER: US/09/631,534

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/092,371

PRIOR FILING DATE: 1998-07-10

PRIOR APPLICATION NUMBER: 60/147, 410

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 09/351,414

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 696

TYPE: PRT

ORGANISM: Homo sapiens

US-09-631-534-2

Query Match 100.0%; Score 630; DB 20; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60

DB 373 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 432

QY 61 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 104

DB 433 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 476

## RESULT 3

US-09-809-617-2  
Sequence 2, Application US/09809617

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Baidur, Nand

APPLICANT: Deisher, Theresa A.

APPLICANT: Bishop, Paul D.

TITLE OF INVENTION: DISINTEGRIN HOMOLOG

FILE REFERENCE: 98-29

CURRENT APPLICATION NUMBER: US/09/809,617

PRIOR FILING DATE: 2001-03-15

PRIOR APPLICATION NUMBER: US/09/351,414

PRIOR FILING DATE: 1999-07-09

NUMBER OF SEQ ID NOS: 13

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 696

TYPE: PRT

ORGANISM: Homo sapiens

US-09-809-617-2

Query Match 100.0%; Score 630; DB 23; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60

DB 373 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 432

QY 61 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 104

DB 433 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 476

## RESULT 4

US-09-809-790-2  
Sequence 2, Application US/09809790

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Baidur, Nand

APPLICANT: Deisher, Theresa A.

APPLICANT: Bishop, Paul D.

TITLE OF INVENTION: DISINTEGRIN HOMOLOG

FILE REFERENCE: 98-29

CURRENT APPLICATION NUMBER: US/09/809,790

PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 09/351,414

PRIOR FILING DATE: 1999-07-09

NUMBER OF SEQ ID NOS: 13

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 696

TYPE: PRT

ORGANISM: Homo sapiens

US-09-809-790-2

Query Match 100.0%; Score 630; DB 23; Length 696;  
Best Local Similarity 100.0%; Pred. No. 6.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60

DB 373 PTKLEPTECGNGYVAGEECDCGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 432

QY 61 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 104

DB 433 RGYECRDVANECDITEYCTGDSGQCPNLAHKODGYACNONOGRG 476

## RESULT 5

US-10-260-506-2  
Sequence 2, Application US/10260506

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Baidur, Nand

APPLICANT: Deisher, Theresa A.

APPLICANT: Bishop, Paul D.

TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP

FILE REFERENCE: 98-29C1

CURRENT APPLICATION NUMBER: US/10/260,506

PRIOR FILING DATE: 2002-09-27

PRIOR APPLICATION NUMBER: 60/092,371

PRIOR FILING DATE: 1998-07-10

PRIOR APPLICATION NUMBER: 60/147, 410

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 09/351,414

PRIOR FILING DATE: 1999-07-09

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 696

TYPE: PRT

ORGANISM: Homo sapiens



TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-260-506-2

Query Match  
Best Local Similarity 100.0%; Score 630; DB 28; Length 696;  
Pred. No. 6.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 60  
DB 373 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 432

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 104  
DB 433 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 475

## RESULT 6

US-09-631-534-10  
Sequence 10, Application US/09631534

GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baindur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
APPLICANT: Talt, David W.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHB  
FILE REFERENCE: 98-28C1  
CURRENT APPLICATION NUMBER: US/09/631,534  
CURRENT FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/092,371  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/147, 410  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 709  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-631-534-10

Query Match  
Best Local Similarity 100.0%; Score 630; DB 20; Length 709;  
Pred. No. 6.7e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHSDPCCNNTSCLFOP 60  
DB 373 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHSDPCCNNTSCLFOP 432

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 104  
DB 433 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 476

## RESULT 7

US-10-260-506-10  
Sequence 10, Application US/10260506

GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baindur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
APPLICANT: Talt, David W.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHB  
FILE REFERENCE: 98-29C1  
CURRENT APPLICATION NUMBER: US/10/260,506  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: 60/092,371  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/147, 410

PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 09/351,414  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 709  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-260-506-10

Query Match  
Best Local Similarity 100.0%; Score 630; DB 28; Length 709;  
Pred. No. 6.7e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 60  
DB 373 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 432

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 104  
DB 433 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 476

## RESULT 8

US-09-760-484-315  
Sequence 315, Application US/09760484

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT243  
CURRENT APPLICATION NUMBER: US/09/760,484  
CURRENT FILING DATE: 2001-01-16  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 638  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 315  
LENGTH: 776  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (731)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-760-484-315

Query Match  
Best Local Similarity 100.0%; Score 630; DB 22; Length 776;  
Pred. No. 7.2e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 60  
DB 440 PTKLFEPTGCGNGVGEAGEECDCGFHVECYGLCKCKSLNSGAHCSDPCCNNTSCLFOP 499

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 104  
DB 500 RGYECRDVAVNECDITEYCTGDSGCCPPNLHKODGYACNONOGRG 543

## RESULT 9

US-09-634-252A-4  
Sequence 4, Application US/09634252A

GENERAL INFORMATION:  
APPLICANT: Cerretti, Douglas P.  
TITLE OF INVENTION: SVPH-15 AND SVPH-17 DNA AND POLYPEPTIDES  
FILE REFERENCE: 03260.0051-00304  
CURRENT APPLICATION NUMBER: US/09/634,252A  
CURRENT FILING DATE: 2000-08-07  
PRIOR APPLICATION NUMBER: 60/074,310  
PRIOR FILING DATE: 1998-02-11  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4

LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-634-252A-4

Query Match 100.0%; Score 630; DB 20; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTCEGNGYVEAGEBCDGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60  
Db 496 PTKLFEPTCEGNGYVEAGEBCDGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 555

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPVLHKODGYACNONGRC 104  
Db 556 RGYECRDVAVNECDITEYCTGDSGCCPVLHKODGYACNONGRC 599

RESULT 10  
US-09-791-537-64795  
Sequence 64795, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 64795  
LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-64795

Query Match 100.0%; Score 630; DB 22; Length 832;  
Best Local Similarity 100.0%; Pred. No. 7.6e-46;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTKLFEPTCEGNGYVEAGEBCDGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 60  
Db 496 PTKLFEPTCEGNGYVEAGEBCDGFHVECYGLCKKCSLSNGAHCSGDPCCNNTSCLFOP 555

QY 61 RGYECRDVAVNECDITEYCTGDSGCCPVLHKODGYACNONGRC 104  
Db 556 RGYECRDVAVNECDITEYCTGDSGCCPVLHKODGYACNONGRC 599

Search completed: October 21, 2003, 10:20:42  
Job time : 316.616 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:08:40 ; Search time 9.33962 Seconds  
(without alignments)  
190.928 Million cell updates/sec

Title: US-09-634-252A-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGAHCSGDPCCNNTSC.....CDITEYCTGSGQCPNLHK 55

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 153893 seqs, 32421815 residues

Total number of hits satisfying chosen parameters: 153893

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patente\_AA New:\*  
1: /cgn2\_6/ptodata/1/paa/PCF\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	832	7	US-60-495-114-1465 Sequence 1465, Ap
2	332	100.0	832	7	US-60-495-114-1466 Sequence 1466, Ap
3	213	64.2	369	7	US-60-495-114-1961 Sequence 1961, Ap
4	213	64.2	623	7	US-60-487-610-2161 Sequence 2161, Ap
5	213	64.2	623	7	US-60-485-450-1482 Sequence 1482, Ap
6	213	64.2	623	7	US-60-495-114-1962 Sequence 1962, Ap
7	213	64.2	769	7	US-60-487-610-2160 Sequence 2160, Ap
8	213	64.2	769	7	US-60-485-450-1481 Sequence 1481, Ap
9	213	64.2	769	7	US-60-495-114-1963 Sequence 1963, Ap
10	213	64.2	773	7	US-60-495-114-1964 Sequence 1964, Ap
11	213	64.2	775	7	US-60-487-610-2162 Sequence 2162, Ap
12	213	64.2	775	7	US-60-485-450-1483 Sequence 1483, Ap
13	185	55.7	761	7	US-60-495-114-1301 Sequence 1301, Ap
14	185	55.7	761	7	US-60-495-114-1302 Sequence 1302, Ap
15	185	55.7	823	7	US-60-490-890-1043 Sequence 1043, Ap
16	185	55.7	823	7	US-60-495-114-1304 Sequence 1304, Ap
17	185	55.7	823	7	US-60-495-114-1309 Sequence 1309, Ap
18	185	55.7	853	7	US-60-495-114-1310 Sequence 1310, Ap
19	185	55.7	853	7	US-60-490-890-1039 Sequence 1039, Ap
20	185	55.7	864	7	US-60-495-114-1308 Sequence 1308, Ap
21	185	55.7	870	7	US-60-490-890-1041 Sequence 1041, Ap
22	185	55.7	870	7	US-60-495-114-1305 Sequence 1305, Ap
23	185	55.7	870	7	US-60-495-114-1311 Sequence 1311, Ap
24	185	55.7	893	7	US-60-495-114-1307 Sequence 1307, Ap
25	185	55.7	899	7	US-60-495-114-1303 Sequence 1303, Ap
26	185	55.7	899	7	US-60-495-114-1306 Sequence 1306, Ap

27	185	55.7	900	7	US-60-495-114-1312 Sequence 1312, Ap
28	168.5	50.8	336	7	US-60-495-114-1584 Sequence 1584, Ap
29	168.5	50.8	398	1	PCT-US03-26780-1448 Sequence 1448, Ap
30	168.5	50.8	409	7	US-60-495-114-1583 Sequence 1583, Ap
31	168.5	50.8	652	7	US-60-495-114-1585 Sequence 1585, Ap
32	168.5	50.8	778	7	US-60-487-610-1926 Sequence 1926, Ap
33	168.5	50.8	781	7	US-60-487-610-1924 Sequence 1924, Ap
34	168.5	50.8	802	7	US-60-487-610-1925 Sequence 1925, Ap
35	168.5	50.8	814	1	PCT-US03-28751-4 Sequence 4, Appl1
36	168.5	50.8	814	7	US-60-490-890-1713 Sequence 1713, Ap
37	167	50.3	499	6	US-10-664-456-6 Sequence 6, Appl1
38	167	50.3	781	6	US-10-664-456-16 Sequence 16, Appl1
39	167	50.3	790	6	US-10-664-456-15 Sequence 15, Appl1
40	167	50.3	790	6	US-10-156-028A-2 Sequence 2, Appl1
41	167	50.3	790	7	US-60-495-114-1565 Sequence 1565, Ap
42	160	48.2	391	1	PCT-US03-30720-2211 Sequence 2211, Ap
43	160	48.2	391	1	PCT-US03-30720-2490 Sequence 2490, Ap
44	160	48.2	746	6	US-10-670-184-4 Sequence 4, Appl1
45	160	48.2	787	1	PCT-US03-30720-951 Sequence 951, Ap

ALIGNMENTS

RESULT 1  
US-60-495-114-1465  
Sequence 1465, Application US/60495114  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
FILE REFERENCE: CLO01480  
CURRENT APPLICATION NUMBER: US/60/495,114  
CURRENT FILING DATE: 2003-08-15  
NUMBER OF SEQ ID NOS: 91238  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1465  
LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-495-114-1465  
Query Match 100.0%; Score 332; DB 7; Length 832;  
Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRAVNECDITEYCTGDSGQCPNLHK 55  
DB 532 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRAVNECDITEYCTGDSGQCPNLHK 586  
RESULT 2  
US-60-495-114-1466  
Sequence 1466, Application US/60495114  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
FILE REFERENCE: CLO01480  
CURRENT APPLICATION NUMBER: US/60/495,114  
CURRENT FILING DATE: 2003-08-15  
NUMBER OF SEQ ID NOS: 91238  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1466  
LENGTH: 832  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-495-114-1466  
Query Match 100.0%; Score 332; DB 7; Length 832;  
Best Local Similarity 100.0%; Pred. No. 5.5e-16;

;; TITLE OF INVENTION:	GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION:	RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
;; TITLE OF INVENTION:	VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES

Query Match	64.2%	Score 213	DB 7	Length 769
Best Local Similarity	65.5%	Pred. No.	5.5e-08	
Matches 36	Conservative	6	Mismatches 11	Indels 2
			Gaps	1

OY 1 CSLSNGAHCSDBGCCNNTSCLFQPRGYECRDVANECDITEYCTGDSGCCPPNLHK 55  
Db 477 CTLLTHDAMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPNLHK 529

RESULT 8  
US-60-485-450-1481

; Sequence 1481, Application US/60485450  
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele  
; APPLICANT: CHANG, Sheng-Yung  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; FILE REFERENCE: CL001470  
; CURRENT APPLICATION NUMBER: US/60/485,450  
; CURRENT FILING DATE: 2003-07-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1481  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-485-450-1481

Query Match 64.2%; Score 213; DB 7; Length 763;  
Best Local Similarity 65.5%; Pred. No. 5.5e-08;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGAHCSDBGCCNNTSCLFQPRGYECRDVANECDITEYCTGDSGCCPPNLHK 55  
Db 477 CTLLTHDAMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPNLHK 529

RESULT 9  
US-60-495-114-1963

; Sequence 1963, Application US/60495114  
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; FILE REFERENCE: CL001480  
; CURRENT APPLICATION NUMBER: US/60/495,114  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1963  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-495-114-1963

Query Match 64.2%; Score 213; DB 7; Length 765;  
Best Local Similarity 65.5%; Pred. No. 5.5e-08;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGAHCSDBGCCNNTSCLFQPRGYECRDVANECDITEYCTGDSGCCPPNLHK 55  
Db 477 CTLLTHDAMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPNLHK 529

RESULT 10  
US-60-495-114-1964

; Sequence 1964, Application US/60495114  
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; FILE REFERENCE: CL001480  
; CURRENT APPLICATION NUMBER: US/60/495,114

; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1964  
; LENGTH: 773  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-495-114-1964

Query Match 64.2%; Score 213; DB 7; Length 773;  
Best Local Similarity 65.5%; Pred. No. 5.5e-08;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGAHCSDBGCCNNTSCLFQPRGYECRDVANECDITEYCTGDSGCCPPNLHK 55  
Db 477 CTLLTHDAMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPNLHK 529

Search completed: October 21, 2003, 10:21:22  
Job time : 9.33962 secs





LOCATION: (323)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-760-484-477

Query Match 100.0%; Score 332; DB 22; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 55  
DB 68 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 122

RESULT 2  
US-10-211-364-1048  
Sequence 1048, Application US/10211364

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P216CIN

CURRENT APPLICATION NUMBER: US/10/211,364

CURRENT FILING DATE: 2002-08-05

PRIOR APPLICATION NUMBER: 09/760,486

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: 60/214,886

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 60/217,487

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,758

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/220,963

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/217,496

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,447

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/218,290

PRIOR FILING DATE: 2000-07-14

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1778

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1048

LENGTH: 368

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (9)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (323)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-211-364-1048

Query Match 100.0%; Score 332; DB 28; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 55  
DB 68 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 122

RESULT 3

US-10-216-893-296  
Sequence 296, Application US/10216893

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P216CIN

CURRENT APPLICATION NUMBER: US/10/216,893

CURRENT FILING DATE: 2002-08-13

PRIOR APPLICATION NUMBER: 09/760,454

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: 60/214,886

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 60/217,487

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,758

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/220,963

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/217,496

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/225,447

PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/218,290

PRIOR FILING DATE: 2000-07-14

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 573

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 296

LENGTH: 368

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (9)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (323)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-216-893-296

Query Match 100.0%; Score 332; DB 28; Length 368;  
Best Local Similarity 100.0%; Pred. No. 5e-25;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 55  
DB 68 CSLSGAHGSDGPCNNNTSCLFQPRGYECRDVAVNECDITEXCTGDSGCCPPLHK 122

RESULT 4  
US-10-217-651-431

Sequence 431, Application US/10217651

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P216CIN

CURRENT APPLICATION NUMBER: US/10/217,651

CURRENT FILING DATE: 2002-08-14

PRIOR APPLICATION NUMBER: 09/760,491

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04



PRIOR APPLICATION NUMBER: 60/214,886  
 PRIOR FILING DATE: 2000-06-28  
 PRIOR APPLICATION NUMBER: 60/217,487  
 PRIOR FILING DATE: 2000-07-11  
 PRIOR APPLICATION NUMBER: 60/225,758  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/220,963  
 PRIOR FILING DATE: 2000-07-26  
 PRIOR APPLICATION NUMBER: 60/217,496  
 PRIOR FILING DATE: 2000-07-11  
 PRIOR APPLICATION NUMBER: 60/225,447  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/218,290  
 PRIOR FILING DATE: 2000-07-14  
 PRIOR APPLICATION NUMBER: 60/225,757  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/226,868  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/216,647  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: 60/225,267  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/216,880  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: 60/225,270  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/251,869  
 PRIOR FILING DATE: 2000-12-08  
 PRIOR APPLICATION NUMBER: 60/235,834  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: 60/234,274  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: 60/234,223  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: 60/228,924  
 PRIOR FILING DATE: 2000-08-30  
 PRIOR APPLICATION NUMBER: 60/224,518  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/236,369  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/224,519  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/220,964  
 PRIOR FILING DATE: 2000-07-26  
 PRIOR APPLICATION NUMBER: 60/241,809  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/249,299  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/236,327  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/241,785  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/244,617  
 PRIOR FILING DATE: 2000-11-01  
 PRIOR APPLICATION NUMBER: 60/225,268  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/236,368  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/251,856  
 PRIOR FILING DATE: 2000-12-08  
 PRIOR APPLICATION NUMBER: 60/251,868  
 PRIOR FILING DATE: 2000-12-08  
 PRIOR APPLICATION NUMBER: 60/229,344  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/234,997  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: 60/229,343  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,345  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,287  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,513

PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/231,413  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/229,509  
 PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/236,367  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/237,039  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,038  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/236,370  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/236,802  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,037  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,040  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/240,960  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/239,935  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/239,937  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/241,787  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,474  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/246,532  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/249,216  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,210  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/226,681  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,759  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/225,213  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/227,182  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,214  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/235,836  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: 60/230,438  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/215,135  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: 60/225,266  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/249,218  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,208  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,213  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,212  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,207  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,245  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,244  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,217  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,211  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,215

```

; OTHER INFORMATION: Fusion Construct
US-09-792-200B-18
Query Match          100.0%; Score 332; DB 28; Length 368;
Best Local Similarity 100.0%; Pred. No. 5e-25;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 55
DB 68 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 122

RESULT 5
US-09-792-200B-18
; Sequence 18, Application US/09792200B
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Fanslow, William C.
; APPLICANT: Poindexter, Kurt M.
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Black, Roy A.
; TITLE OF INVENTION: INTEGRIN ANTAGONISTS
; FILE REFERENCE: 2958-A
; CURRENT APPLICATION NUMBER: US/09/792,200B
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/184,865
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: Fusion Construct
US-09-792-200C-18
Query Match          100.0%; Score 332; DB 22; Length 540;
Best Local Similarity 100.0%; Pred. No. 6.9e-25;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 55
DB 50 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 104

RESULT 6
US-09-792-200C-18
; Sequence 18, Application US/09792200C
; GENERAL INFORMATION:
; APPLICANT: Fanslow III, William C.
; APPLICANT: Poindexter, Kurt M.
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Black, Roy A.
; TITLE OF INVENTION: INTEGRIN ANTAGONISTS
; FILE REFERENCE: 2958-A
; CURRENT APPLICATION NUMBER: US/09/792,200C
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/184,865
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion Construct
US-09-792-200C-18
Query Match          100.0%; Score 332; DB 22; Length 540;
Best Local Similarity 100.0%; Pred. No. 6.9e-25;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 55
DB 50 CSLSNAGHCSGDPCCNNTSCIFQPRGYECRDVAVECDITTEYCTGDSGCCPPNHLK 104

RESULT 7
US-09-949-016-11203
; Sequence 11203, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11203
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11203
Query Match          100.0%; Score 332; DB 24; Length 686;
Best Local Similarity 100.0%; Pred. No. 8.4e-25;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 55  
 Db 386 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 440

RESULT 8

US-09-631-534-2  
 ; Sequence 2, Application US/09631534  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; APPLICANT: Baindur, Nand  
 ; APPLICANT: Deisher, Theresa A.  
 ; APPLICANT: Bishop, Paul D.  
 ; APPLICANT: Tait, David W.  
 ; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP  
 ; FILE REFERENCE: 98-29C1  
 ; CURRENT APPLICATION NUMBER: US/09/631,534  
 ; CURRENT FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: 60/092,371  
 ; PRIOR FILING DATE: 1998-07-10  
 ; PRIOR APPLICATION NUMBER: 60/147, 410  
 ; PRIOR FILING DATE: 1999-08-05  
 ; PRIOR APPLICATION NUMBER: 09/351,414  
 ; PRIOR FILING DATE: 1999-07-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 696  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-631-534-2

Query Match 100.0%; Score 332; DB 20; Length 696;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-25;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 55  
 Db 409 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 463

RESULT 9

US-09-809-617-2  
 ; Sequence 2, Application US/09809617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; APPLICANT: Baindur, Nand  
 ; APPLICANT: Deisher, Theresa A.  
 ; APPLICANT: Bishop, Paul D.  
 ; TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
 ; FILE REFERENCE: 98-29  
 ; CURRENT APPLICATION NUMBER: US/09/809,617  
 ; CURRENT FILING DATE: 2001-03-15  
 ; PRIOR APPLICATION NUMBER: US/09/351,414  
 ; PRIOR FILING DATE: 1999-07-09  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 696  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-809-617-2

Query Match 100.0%; Score 332; DB 23; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-25;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 55  
 Db 409 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 463

RESULT 10

US-09-809-790-2  
 ; Sequence 2, Application US/09809790  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; APPLICANT: Baindur, Nand  
 ; APPLICANT: Deisher, Theresa A.  
 ; APPLICANT: Bishop, Paul D.  
 ; TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
 ; FILE REFERENCE: 98-29  
 ; CURRENT APPLICATION NUMBER: US/09/809,790  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 09/351,414  
 ; PRIOR FILING DATE: 1999-07-09  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 696  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-809-790-2

Query Match 100.0%; Score 332; DB 23; Length 696;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-25;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 55  
 Db 409 CSLSNGAHCSDGPCNNNTSCLFQPRGYECRDAVNECDITTEYCTGDSGCCPPNLRK 463

Search completed: October 21, 2003, 10:20:42  
 Job time : 166.384 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:27:07 ; Search time 1871.38 Seconds

(without alignments)  
974.684 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586

Percent score: 332  
Sequence: 1 CSUSNGAHCSDGPCNNNTSC.....CDITEYCTGDSGQCPNHLHK 55

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 33363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US0964222/runat_21102003_090326_3436/app_query.fasca_1.462
-DB=Pending_Patents_NA_Main -QFMT=fastcap -SUFFIX=p2n.rnp -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum2
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09634252 @CGN 1.1 5064 @runat_21102003_090326_3436
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPHLOCK=100
-LONLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: Pending Patents_NA_Main:*
2: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.oid:*
4: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
```

```
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq:*
```

102: /cgn2\_6/prodata/2/pna/US04\_COMB.seq.\*

## SUMMARIES

Result		Query				Description	
No.	Score	Match	Length	DB	ID		
1	332	100.0	1668	33	US-09-792-200B-17	Sequence 17, Appl	
2	332	100.0	1668	33	US-09-792-200C-17	Sequence 17, Appl	
3	332	100.0	2268	27	US-09-631-534-1	Sequence 1, Appl	
4	332	100.0	2268	34	US-09-809-617-1	Sequence 1, Appl	
5	332	100.0	2268	34	US-09-809-790-1	Sequence 1, Appl	
6	332	100.0	2268	49	US-10-960-506-1	Sequence 1, Appl	
7	332	100.0	2394	40	US-09-949-016-5332	Sequence 5332, Appl	
8	332	100.0	2499	27	US-09-634-252A-2	Sequence 2, Appl	
9	332	100.0	2899	47	US-10-202-675-2	Sequence 2, Appl	
10	332	100.0	2899	47	US-10-170-235-22436	Sequence 22436, Appl	
11	332	100.0	2959	21	US-09-496-321-82	Sequence 82, Appl	
12	332	100.0	2959	21	US-09-496-321-109	Sequence 109, Appl	
13	332	100.0	2959	65	US-60-118-905-82	Sequence 82, Appl	
14	332	100.0	2959	65	US-60-118-905-109	Sequence 109, Appl	
15	332	100.0	3054	46	US-10-126-052A-149	Sequence 149, Appl	
16	332	100.0	3054	46	US-10-126-052A-124	Sequence 134, Appl	
17	332	100.0	3402	33	US-09-760-484-117	Sequence 117, Appl	
18	332	100.0	3402	48	US-10-211-364-395	Sequence 395, Appl	
19	332	100.0	3402	48	US-10-216-893-114	Sequence 114, Appl	
20	332	100.0	3402	34	US-10-217-651-164	Sequence 164, Appl	
21	332	100.0	4043	38	US-09-824-129-2	Sequence 2, Appl	
22	332	100.0	4043	51	US-10-359-464-12	Sequence 2, Appl	
23	332	100.0	4440	33	US-09-760-484-15	Sequence 15, Appl	
24	332	100.0	4720	47	US-09-631-534-9	Sequence 9, Appl	
25	332	100.0	4720	29	US-10-260-506-9	Sequence 29, Appl	
26	332	100.0	6391	71	US-60-172-360-21719	Sequence 21719, Appl	
27	332	100.0	6400	42	US-09-976-594-61	Sequence 61, Appl	
28	332	100.0	6400	79	US-60-240-409-61	Sequence 61, Appl	
29	332	100.0	6400	87	US-60-324-185-1573	Sequence 1573, Appl	
30	325	97.9	494	17	US-09-234-611-16347	Sequence 16347, Appl	
31	325	97.9	494	17	US-09-234-611-16347	Sequence 22101, Appl	
32	325	97.9	494	18	US-09-235-076-22101	Sequence 22101, Appl	
33	325	97.9	494	18	US-09-289-766-20173	Sequence 20173, Appl	
34	325	97.9	494	19	US-09-332-782-22101	Sequence 22101, Appl	
35	325	97.9	494	32	US-09-737-221-22101	Sequence 22101, Appl	
36	325	97.9	494	38	US-09-904-809-16347	Sequence 16347, Appl	
37	325	97.9	494	38	US-09-918-995-22101	Sequence 22101, Appl	
38	319	96.1	2271	46	US-10-144-771-15737	Sequence 20173, Appl	
39	319	96.1	2271	91	US-60-360-207-15737	Sequence 15737, Appl	
40	319	96.1	6401	75	US-60-213-359-127	Sequence 127, Appl	
41	315	94.9	468	22	US-09-519-705-803	Sequence 803, Appl	
42	315	94.9	468	25	US-09-574-454-803	Sequence 803, Appl	
43	315	94.9	468	48	US-10-221-279-792	Sequence 792, Appl	
44	308	92.8	261	7	US-08-275-627A-52	Sequence 52, Appl	
45	308	92.8	261	7	US-08-275-627A-52	Sequence 52, Appl	

```
US-09-634-252a-4_COPY_532_586 (1-55) x US-09-792-200C-17 (1-1668)
Qy 1 CySerleuSeranglYAlaHisCySerAspGlyProCyScySaenAnThSerCys 20
Db 172 TGTTCCTCTCCAAAGGGGCTCACTGACGACGCGGCTCTGTAACAATACCTCATGT 231
Qy 21 LeuPheGlnProArgGlyTyrgluCyArgAspAlaValansgluCyAspIleThrglu 40
Db 232 CTTTTCAGCCACGAGGATGTAATGCGGGATGCTGTGAACGAGTGTGATTACTGAA 291
Qy 41 TyrCyThrGlyAspSerGlyGlnCyProProAsnLeuHislys 55
Db 292 TATTGTACTGAGACTCTGTGTGATGCGCCACCAATCTTCATAAG 336

RESULT 3
US-09-631-534-1
; Sequence 1, Application US/09631534
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Telf, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/09/631,534
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2090)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2268)
; OTHER INFORMATION: n = A,T,C or G
US-09-631-534-1

Alignment Scores:
Pred. No.: 3,79e-28 Length: 2268
Percent Similarity: 332.00 Matches: 55
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
Indels: 0
Gaps: 27
DB:

US-09-634-252a-4_COPY_532_586 (1-55) x US-09-631-534-1 (1-2268)
Qy 1 CySerleuSeranglYAlaHisCySerAspGlyProCyScySaenAnThSerCys 20
Db 1227 TGTTCCTCTCCAAAGGGGCTCACTGACGACGCGGCTCTGTAACAATACCTCATGT 1286
Qy 21 LeuPheGlnProArgGlyTyrgluCyArgAspAlaValansgluCyAspIleThrglu 40
Db 1287 CTTTTCAGCCACGAGGATGTAATGCGGGATGCTGTGAACGAGTGTGATTACTGAA 1346
Qy 41 TyrCyThrGlyAspSerGlyGlnCyProProAsnLeuHislys 55
Db 1347 TATTGTACTGAGACTCTGTGTGATGCGCCACCAATCTTCATAAG 1391

RESULT 4
US-09-809-617-1
; Sequence 1, Application US/09809617
; GENERAL INFORMATION:
```

```
APPLICANT: Sheppard, Paul O.
APPLICANT: Baidur, Nand
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: DISINTEGRIN HOMOLOG
FILE REFERENCE: 98-29
CURRENT APPLICATION NUMBER: US/09/809,617
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: US/09/351,414
PRIOR FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2268
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(2090)
NAME/KEY: misc_feature
LOCATION: (1)...(2268)
OTHER INFORMATION: n = A,T,C or G
US-09-809-790-1

Alignment Scores:
Pred. No.: 3,79e-28 Length: 2268
Percent Similarity: 332.00 Matches: 55
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
Indels: 0
Gaps: 34
DB:

US-09-634-252a-4_COPY_532_586 (1-55) x US-09-809-617-1 (1-2268)
Qy 1 CySerleuSeranglYAlaHisCySerAspGlyProCyScySaenAnThSerCys 20
Db 1227 TGTTCCTCTCCAAAGGGGCTCACTGACGACGCGGCTCTGTAACAATACCTCATGT 1286
Qy 21 LeuPheGlnProArgGlyTyrgluCyArgAspAlaValansgluCyAspIleThrglu 40
Db 1287 CTTTTCAGCCACGAGGATGTAATGCGGGATGCTGTGAACGAGTGTGATTACTGAA 1346
Qy 41 TyrCyThrGlyAspSerGlyGlnCyProProAsnLeuHislys 55
Db 1347 TATTGTACTGAGACTCTGTGTGATGCGCCACCAATCTTCATAAG 1391

RESULT 5
US-09-809-790-1
; Sequence 1, Application US/09809790
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/809,790
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2090)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2268)
; OTHER INFORMATION: n = A,T,C or G
US-09-809-790-1
```

## Alignment Scores:

Pred. No.: 3,798-28 Length: 2268  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 34 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-09-809-790-1 (1-2268)

QY 1 CysSerLeuSerAseGlyAlaHisCysSerAseGlyProCysCysAseAseHisSerCys 20

Db 1227 TGTTCCTCTCCACAGGGGCTCAGTCAGCGAGCGGGCCCTGCTGTAAACAATCCTCATGT 1286

QY 21 LeuPheGlnProAseGlyTyrGluCysAseAseAlaValaAsnGluCysAsePileThrglu 40

Db 1287 CTTTTCAGCCAGAGGGATGAATGCCGGATGCTGTGAACGAGTGTGATTTACTGAA 1346

QY 41 TyrCysThrglyAseSerGlyGlnCysProProAseLeuHisLys 55

Db 1347 TATTGTACTGAGACTCTGTGTCAGTCCCAACAATCTTCAATAG 1391

## RESULT 6

US-10-260-506-1

Sequence 1, Application US/10260506

GENERAL INFORMATION:

APPLICANT: Shepard, Paul O.

APPLICANT: Baird, Nand

APPLICANT: Deisher, Theresa A.

APPLICANT: Bishop, Paul D.

APPLICANT: Taft, David W.

TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP

FILE REFERENCE: 98-29C1

CURRENT APPLICATION NUMBER: US/10/260,506

CURRENT FILING DATE: 2002-09-27

PRIOR APPLICATION NUMBER: 60/092,371

PRIOR FILING DATE: 1998-07-10

PRIOR APPLICATION NUMBER: 60/147, 410

PRIOR FILING DATE: 1999-08-05

PRIOR APPLICATION NUMBER: 09/351,414

PRIOR FILING DATE: 1999-07-09

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 2268

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (3)...(2090)

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(2268)

OTHER INFORMATION: n = A,T,C or G

US-10-260-506-1

## Alignment Scores:

Pred. No.: 3,798-28 Length: 2268  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 49 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-10-260-506-1 (1-2268)

QY 1 CysSerLeuSerAseGlyAlaHisCysSerAseGlyProCysCysAseAseHisSerCys 20

Db 1227 TGTTCCTCTCCACAGGGGCTCAGTCAGCGAGCGGGCCCTGCTGTAAACAATCCTCATGT 1286

QY 21 LeuPheGlnProAseGlyTyrGluCysAseAseAlaValaAsnGluCysAsePileThrglu 40

Db 1287 CTTTTCAGCCAGAGGGATGAATGCCGGATGCTGTGAACGAGTGTGATTTACTGAA 1346

QY 41 TyrCysThrglyAseSerGlyGlnCysProProAseLeuHisLys 55

Db 1347 TATTGTACTGAGACTCTGTGTCAGTCCCAACAATCTTCAATAG 1391

## RESULT 7

US-09-949-016-5332

Sequence 5332, Application US/09949016

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5332

LENGTH: 2394

TYPE: DNA

ORGANISM: Human

US-09-949-016-5332

## Alignment Scores:

Pred. No.: 4,046-28 Length: 2394  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 40 Gaps: 0

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-09-949-016-5332 (1-2394)

QY 1 CysSerLeuSerAseGlyAlaHisCysSerAseGlyProCysCysAseAseHisSerCys 20

Db 1158 TGTTCCTCTCCACAGGGGCTCAGTCAGCGAGCGGGCCCTGCTGTAAACAATCCTCATGT 1217

QY 21 LeuPheGlnProAseGlyTyrGluCysAseAseAlaValaAsnGluCysAsePileThrglu 40

Db 1218 CTTTTCAGCCAGAGGGATGAATGCCGGATGCTGTGAACGAGTGTGATTTACTGAA 1277

QY 41 TyrCysThrglyAseSerGlyGlnCysProProAseLeuHisLys 55

Db 1278 TATTGTACTGAGACTCTGTGTCAGTCCCAACAATCTTCAATAG 1322

## RESULT 8

US-09-634-252a-2

Sequence 2, Application US/09634252A

GENERAL INFORMATION:

APPLICANT: Cerretti, Douglas P.

TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES

FILE REFERENCE: 03260, 0051-00304

CURRENT APPLICATION NUMBER: US/09/634, 252A

CURRENT FILING DATE: 2000-08-07

PRIOR APPLICATION NUMBER: 60/074,310

PRIOR FILING DATE: 1998-02-11

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 2499

TYPE: DNA

ORGANISM: Homo sapiens

US-09-634-252a-2

## Alignment Scores:

Pred. No.: 4,246-28 Length: 2499



Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 27 Gaps: 0

US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-09-634-252A-2 (1-2499)

QY 1 CysSerLeuSerAenGlyAlaHisCysSerAaspGlyProCysAaAnaenThSerCys 20  
Db 1594 TGTTCCTCTCCAAAGGGGCTCACTGACGACGAGGGGCTCTGTAACTAATCTCATGT 1653  
QY 21 LeuPheGlnProAArgGlyTyrgluCysArgAaspAlaValaenGluCysAaspIleThrglu 40  
Db 1654 CTTTTCAGCCACAGAGGTATGAATGCCGGGATGCTGTGAACGATGTGATATTACTGAA 1713  
QY 41 TyrCysThrGlyAaspSerGlyGlnCysProProAasnLeuHisIlys 55  
Db 1714 TATTGTACTGAGACTGTGTGACGTGCCCAATCTTCATAAG 1758

## RESULT 9

US-10-202-675-2  
; Sequence 2, Application US/10202675  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: SVP3-13 AND SVP3-17 DNA AND POLYPEPTIDES  
; FILE REFERENCE: 03260.0051-00304  
; CURRENT APPLICATION NUMBER: US/10/202,675  
; CURRENT FILING DATE: 2002-07-23  
; PRIOR APPLICATION NUMBER: US/09/634,252  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: 60/074,310  
; PRIOR FILING DATE: 1998-02-11  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2499  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-202-675-2

Alignment Scores:  
Pred. No.: 4,24e-28 Length: 2499  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 48 Gaps: 0

US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-10-202-675-2 (1-2499)

QY 1 CysSerLeuSerAenGlyAlaHisCysSerAaspGlyProCysAaAnaenThSerCys 20  
Db 1594 TGTTCCTCTCCAAAGGGGCTCACTGACGACGAGGGGCTCTGTAACTAATCTCATGT 1653  
QY 21 LeuPheGlnProAArgGlyTyrgluCysArgAaspAlaValaenGluCysAaspIleThrglu 40  
Db 1654 CTTTTCAGCCACAGAGGTATGAATGCCGGGATGCTGTGAACGATGTGATATTACTGAA 1713  
QY 41 TyrCysThrGlyAaspSerGlyGlnCysProProAasnLeuHisIlys 55  
Db 1714 TATTGTACTGAGACTGTGTGACGTGCCCAATCTTCATAAG 1758

## RESULT 10

US-10-170-235-22436  
; Sequence 22436, Application US/10170235  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig  
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN  
; FILE REFERENCE: C0001380  
; CURRENT APPLICATION NUMBER: US/10/170,235  
; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514  
; SEQ ID NO 22436  
; LENGTH: 2899  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-170-235-22436

Alignment Scores:  
Pred. No.: 5.03e-28 Length: 2899  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 47 Gaps: 0

US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-10-170-235-22436 (1-2899)

QY 1 CysSerLeuSerAenGlyAlaHisCysSerAaspGlyProCysAaAnaenThSerCys 20  
Db 1663 TGTTCCTCTCCAAAGGGGCTCACTGACGACGAGGGGCTCTGTAACTAATCTCATGT 1722  
QY 21 LeuPheGlnProAArgGlyTyrgluCysArgAaspAlaValaenGluCysAaspIleThrglu 40  
Db 1723 CTTTTCAGCCACAGAGGTATGAATGCCGGGATGCTGTGAACGATGTGATATTACTGAA 1782  
QY 41 TyrCysThrGlyAaspSerGlyGlnCysProProAasnLeuHisIlys 55  
Db 1783 TATTGTACTGAGACTGTGTGACGTGCCCAATCTTCATAAG 1827

Search completed: October 21, 2003, 14:37:22  
Job time: 1876.38 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:34:26 ; Search time 110.346 Seconds  
(without alignments)  
758.947 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGAHCSDGPCNNNTSC.....CDITEYCTGDSGQCPEMLHK 55

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1640717 seqs, 761333381 residues

Total number of hits satisfying chosen parameters: 3631434

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cg2n\_1/USPTO.spool/US09634252/runat\_21102003\_090326\_3471/app.query.fasta\_1.462  
-DB=Pending\_Patents\_NA\_New -OFMT=fastcap -SUFFIX=p2n.rnpn -MINMATCH=0.1  
-LOOPEXT=0 -UNIT5=Dits -START=1 -END=1 -MATRIX=D10sum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09634252 @CGN 1 1 248 @runat 21.0203 090326\_3471  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEBUERY -NEG\_SCORES=0 -WAIT -HSEPLECK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELIXT=7

Database : Pending Patents\_NA\_New:\*  
1: /cg2n\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
2: /cg2n\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cg2n\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
4: /cg2n\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cg2n\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cg2n\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
7: /cg2n\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	3760	7	US-60-495-114-329
2	332	100.0	4071	7	US-60-495-114-330
3	213	64.2	1911	7	US-60-487-610-948
4	213	64.2	1911	7	US-60-485-450-577
5	213	64.2	3117	7	US-60-485-114-825
6	213	64.2	3167	7	US-60-487-610-947
7	213	64.2	3167	7	US-60-485-450-576
8	213	64.2	3185	7	US-60-487-610-949
9	213	64.2	3185	7	US-60-485-450-578
10	213	64.2	3421	7	US-60-495-114-827
11	213	64.2	3431	7	US-60-495-114-826

12	213	64.2	3433	7	US-60-495-114-828	Sequence 828, App
13	185	55.7	2636	7	US-60-490-890-1042	Sequence 1042, Ap
14	185	55.7	2744	7	US-60-490-890-1038	Sequence 1038, Ap
15	185	55.7	2775	7	US-60-495-114-168	Sequence 168, App
16	185	55.7	2775	7	US-60-495-114-173	Sequence 173, App
17	185	55.7	2865	7	US-60-495-114-174	Sequence 174, App
18	185	55.7	2876	7	US-60-495-114-166	Sequence 166, App
19	185	55.7	3347	7	US-60-490-890-1040	Sequence 1040, Ap
20	185	55.7	3759	7	US-60-495-114-172	Sequence 172, App
21	185	55.7	3777	7	US-60-495-114-169	Sequence 169, App
22	185	55.7	3777	7	US-60-495-114-175	Sequence 175, App
23	185	55.7	3801	7	US-60-495-114-171	Sequence 171, App
24	185	55.7	3819	7	US-60-495-114-167	Sequence 167, App
25	185	55.7	3819	7	US-60-495-114-170	Sequence 170, App
26	185	55.7	3867	7	US-60-495-114-176	Sequence 176, App
27	185	55.7	3878	7	US-60-495-114-165	Sequence 165, App
28	173.5	52.3	2815	1	PCT-US03-28751-6	Sequence 6, Appl1
29	168.5	50.8	1194	1	PCT-US03-26780-217	Sequence 217, App
30	168.5	50.8	2286	7	US-60-495-114-447	Sequence 447, App
31	168.5	50.8	2313	7	US-60-495-114-448	Sequence 448, App
32	168.5	50.8	2336	7	US-60-495-114-449	Sequence 449, App
33	168.5	50.8	2712	7	US-60-487-610-511	Sequence 511, App
34	168.5	50.8	2740	1	PCT-US03-28751-3	Sequence 3, Appl1
35	168.5	50.8	2740	7	US-60-490-890-1712	Sequence 1712, Ap
36	168.5	50.8	2762	7	US-60-487-610-512	Sequence 512, App
37	168.5	50.8	2808	7	US-60-487-610-513	Sequence 513, App
38	168.5	50.8	4547	7	US-60-495-114-446	Sequence 446, App
39	167	50.3	1500	6	US-10-664-456-3	Sequence 3, Appl1
40	167	50.3	2346	6	US-10-664-456-11	Sequence 11, Appl1
41	167	50.3	2373	6	US-10-664-456-10	Sequence 10, Appl1
42	167	50.3	2373	6	US-10-156-028A-1	Sequence 1, Appl1
43	167	50.3	2536	7	US-60-495-114-429	Sequence 429, App
44	167	50.3	14966	7	US-60-495-114-16886	Sequence 16886, A
45	167	50.3	50248	7	US-60-495-114-16422	Sequence 16422, A

## ALIGNMENTS

RESULT 1  
US-60-495-114-329  
; Sequence 329, Application US/60495114  
; GENERAL INFORMATION:  
; APPLICANT: CARCILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CU001480  
; CURRENT APPLICATION NUMBER: US/60/495,114  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 329  
; LENGTH: 3760  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-495-114-329  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-60-495-114-329 (1-3760)  
Alignment Scores:  
Pred. No.: 2.67e-35 Length: 3760  
Score: 332.00 Matches: 55  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0  
QY 1 CysSerLeuSerArgnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20  
DB 1817 TGTTCCTTCCTCAACGGGCTCACTGCGACGACGAGCCCTGCTTAACAATACCTCATGT 1876  
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspAlaThrGlu 40

```
Db      1877 CTTTTCACCCACGAGGTATTAATGCCGGATGCTGTGACAGAGTGTGATTAATTAAGTAA 1936
Qy      41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db      1937 TATTGTACTGGAGACTCTGTGTCAAGTCCACCAAAATCTTCAATAG 1981

RESULT 2
US-60-495-114-330
; Sequence 330, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 4071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-495-114-330

Alignment Scores:
Pred. No.:      2,96e-35      Length:      4071
Score:          332.00      Matches:      55
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              7          Gaps:          0

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-495-114-330 (1-4071)

Qy      1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20
Db      1817 TGTTCCCTCTCCACGCGGCTCATGTCAGCAGCGGGCCCTGTGTAACAAATACCTCATGT 1876

Qy      21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db      1877 CTTTTCACCCACGAGGGTGTGTAATGCCGGATGCTGTAAAGAGTGTATTTACTGAA 1936

Qy      41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db      1937 TATTGTACTGGAGACTCTGTGTCAAGTCCACCAAAATCTTCAATAG 1981

RESULT 3
US-60-487-610-948
; Sequence 948, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 948
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-948

Alignment Scores:
Pred. No.:      2,93e-19      Length:      1911
Score:          213.00      Matches:      36
Percent Similarity: 76.36%      Conservative: 6
Best Local Similarity: 65.45%      Mismatches: 11
Query Match:    64.16%      Indels:      2
DB:              7          Gaps:          1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-487-610-948 (1-1911)

Qy      1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20
Db      1429 TGCACCTGTACTCAGACGACGCCATGTGCGACGCGGCTCTGTGCGCCG-----TGC 1482

Qy      21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db      1483 AAGTACGAACCAACGCGGGTGTGTCTGCGGAGAGCGCGTGAACGAGTGCACATCCGGAG 1542

Qy      41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db      1543 ACCTCACCAGGGGACTTACGACAGTCCCGCTTAACCTGCACAAG 1587

RESULT 5
US-60-495-114-825
; Sequence 825, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 825
```

```
DB:      7          Gaps:      1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-487-610-948 (1-1911)

Qy      1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20
Db      1429 TGCACCTGTACTCAGACGACGCCATGTGCGACGCGGCTCTGTGCGCCG-----TGC 1482

Qy      21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db      1483 AAGTACGAACCAACGCGGGTGTGTCTGCGGAGAGCGCGTGAACGAGTGCACATCCGGAG 1542

Qy      41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db      1543 ACCTCACCAGGGGACTTACGACAGTCCCGCTTAACCTGCACAAG 1587

RESULT 4
US-60-485-450-577
; Sequence 577, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-577

Alignment Scores:
Pred. No.:      2,93e-19      Length:      1911
Score:          213.00      Matches:      36
Percent Similarity: 76.36%      Conservative: 6
Best Local Similarity: 65.45%      Mismatches: 11
Query Match:    64.16%      Indels:      2
DB:              7          Gaps:          1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-485-450-577 (1-1911)

Qy      1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20
Db      1429 TGCACCTGTACTCAGACGACGCCATGTGCGACGCGGCTCTGTGCGCCG-----TGC 1482

Qy      21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db      1483 AAGTACGAACCAACGCGGGTGTGTCTGCGGAGAGCGCGTGAACGAGTGCACATCCGGAG 1542

Qy      41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
Db      1543 ACCTCACCAGGGGACTTACGACAGTCCCGCTTAACCTGCACAAG 1587
```

```
; LENGTH: 3117
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-495-114-825

Alignment Scores:
Pred. No.: 5.56e-19 Length: 3117
Score: 213.00 Matches: 36
Percent Similarity: 76.36% Conservative: 6
Best Local Similarity: 65.45% Mismatches: 11
Query Match: 64.16% Indels: 2
DB: 7 Gaps: 1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-495-114-825 (1-3117)
OY 1 CysSerLeuSerAnGlyAlaHisCysSerApsGlyProCysCysAaAnAntHrSerCys 20
Db 1115 TGCACCTGACTACGACGCCATGTGACGACGCGGCTGTGTCGCCCGC-----TGC 1168

OY 21 LeuPheGlnProArgGlyTyrgluCysArgAspAlaValaAnGluCysAspIleThrGlu 40
Db 1169 AAGTACGACACCGGGGTGTCTCTGCGGAGAGCCCGTGAACGAGTCGCACATCGCGGAG 1228

OY 41 TyTCysThrGlyApsSerGlyGlnCysProProAnLeuHisIlys 55
Db 1229 ACCTGACCGGGGACTCTAGCCAGTCCCGCTTAACCTGCACAAAG 1273

RESULT 6
US-60-487-610-947
; Sequence 947, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 947
; LENGTH: 3167
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-947

Alignment Scores:
Pred. No.: 5.68e-19 Length: 3167
Score: 213.00 Matches: 36
Percent Similarity: 76.36% Conservative: 6
Best Local Similarity: 65.45% Mismatches: 11
Query Match: 64.16% Indels: 2
DB: 7 Gaps: 1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-487-610-947 (1-3167)
OY 1 CysSerLeuSerAnGlyAlaHisCysSerApsGlyProCysCysAaAnAntHrSerCys 20
Db 1429 TGCACCTGACTACGACGCCATGTGACGACGCGGCTGTGTCGCCCGC-----TGC 1482

OY 21 LeuPheGlnProArgGlyTyrgluCysArgAspAlaValaAnGluCysAspIleThrGlu 40
Db 1483 AAGTACGACACCGGGGTGTCTCTGCGGAGAGCCCGTGAACGAGTCGCACATCGCGGAG 1542

OY 41 TyTCysThrGlyApsSerGlyGlnCysProProAnLeuHisIlys 55
Db 1543 ACCTGACCGGGGACTCTAGCCAGTCCCGCTTAACCTGCACAAAG 1567

RESULT 7
US-60-485-450-576
; Sequence 576, Application US/60485450
; GENERAL INFORMATION:
```

```
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 576
; LENGTH: 3167
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-485-450-576

Alignment Scores:
Pred. No.: 5.68e-19 Length: 3167
Score: 213.00 Matches: 36
Percent Similarity: 76.36% Conservative: 6
Best Local Similarity: 65.45% Mismatches: 11
Query Match: 64.16% Indels: 2
DB: 7 Gaps: 1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-485-450-576 (1-3167)
OY 1 CysSerLeuSerAnGlyAlaHisCysSerApsGlyProCysCysAaAnAntHrSerCys 20
Db 1429 TGCACCTGACTACGACGCCATGTGACGACGCGGCTGTGTCGCCCGC-----TGC 1482

OY 21 LeuPheGlnProArgGlyTyrgluCysArgAspAlaValaAnGluCysAspIleThrGlu 40
Db 1483 AAGTACGACACCGGGGTGTCTCTGCGGAGAGCCCGTGAACGAGTCGCACATCGCGGAG 1542

OY 41 TyTCysThrGlyApsSerGlyGlnCysProProAnLeuHisIlys 55
Db 1543 ACCTGACCGGGGACTCTAGCCAGTCCCGCTTAACCTGCACAAAG 1587

RESULT 8
US-60-487-610-949
; Sequence 949, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487, 610
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 949
; LENGTH: 3185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-487-610-949

Alignment Scores:
Pred. No.: 5.72e-19 Length: 3185
Score: 213.00 Matches: 36
Percent Similarity: 76.36% Conservative: 6
Best Local Similarity: 65.45% Mismatches: 11
Query Match: 64.16% Indels: 2
DB: 7 Gaps: 1

US-09-634-252a-4_COPY_532_586 (1-55) x US-60-487-610-949 (1-3185)
OY 1 CysSerLeuSerAnGlyAlaHisCysSerApsGlyProCysCysAaAnAntHrSerCys 20
Db 1429 TGCACCTGACTACGACGCCATGTGACGACGCGGCTGTGTCGCCCGC-----TGC 1482
```

QY 21 LeupheginProA:rgGlyTyrgluCySa:rgAspaIaValaIsngluCySa:spIleThrglu 40  
Db 1483 AAGTACGAACCGGGGTGTGTCTGCTCCGAGAGCGCGTGAAGAGTGGACATCGCGAG 1542  
QY 41 TyrCysThrgIyAspSerGlyGlnCySa:ProProAsnLeuHisLys 55  
Db 1543 ACTGACCGGGGACTTAAGCCAGTGCCTTAACCTGCACAG 1587

## RESULT 9

US-60-485-450-578  
; Sequence 578, Application US/60485450  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; FILE REFERENCE: CLO01470  
; CURRENT APPLICATION NUMBER: US/60/485,450  
; CURRENT FILING DATE: 2003-07-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 578  
; LENGTH: 3185  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-485-450-578

## Alignment Scores:

Pred. No.: 5.72e-19 Length: 3185  
Score: 213.00 Matches: 36  
Percent Similarity: 76.36% Conservative: 6  
Best Local Similarity: 65.45% Mismatches: 11  
Query Match: 64.16% Indels: 2  
DB: 7 Gaps: 1

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-60-485-450-578 (1-3185)

QY 1 CysSerLeuSerAsnGlyAlaHisCySerAspGlyProCySa:AsnAnThrSerCys 20  
Db 1429 TGCACCTGACTCAGACGACCATGTGACAGCGAGCGCTCTGCTGCGCCG----TGC 1482  
QY 21 LeupheginProA:rgGlyTyrgluCySa:rgAspaIaValaIsngluCySa:spIleThrglu 40  
Db 1483 AAGTACGAACCGGGGTGTGTCTGCTCCGAGAGCGCGTGAAGAGTGGACATCGCGAG 1542  
QY 41 TyrCysThrgIyAspSerGlyGlnCySa:ProProAsnLeuHisLys 55  
Db 1543 ACTGACCGGGGACTTAAGCCAGTGCCTTAACCTGCACAG 1587

## RESULT 10

US-60-495-114-827  
; Sequence 827, Application US/60495114  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: CLO01480  
; CURRENT APPLICATION NUMBER: US/60/495,114  
; CURRENT FILING DATE: 2003-08-15  
; NUMBER OF SEQ ID NOS: 91238  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 827  
; LENGTH: 3421  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-495-114-827

## Alignment Scores:

Pred. No.: 6.28e-19 Length: 3421  
Score: 213.00 Matches: 36

Percent Similarity: 76.36% Conservative: 6  
Best Local Similarity: 65.45% Mismatches: 11  
Query Match: 64.16% Indels: 2  
DB: 7 Gaps: 1

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-60-495-114-827 (1-3421)

QY 1 CysSerLeuSerAsnGlyAlaHisCySerAspGlyProCySa:AsnAnThrSerCys 20  
Db 1429 TGCACCTGACTCAGACGACCATGTGACAGCGAGCGCTCTGCTGCGCCG----TGC 1482  
QY 21 LeupheginProA:rgGlyTyrgluCySa:rgAspaIaValaIsngluCySa:spIleThrglu 40  
Db 1483 AAGTACGAACCGGGGTGTGTCTGCTCCGAGAGCGCGTGAAGAGTGGACATCGCGAG 1542  
QY 41 TyrCysThrgIyAspSerGlyGlnCySa:ProProAsnLeuHisLys 55  
Db 1543 ACTGACCGGGGACTTAAGCCAGTGCCTTAACCTGCACAG 1587

Search completed: October 21, 2003, 14:42:56  
Job time : 115.346 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 12:02:06 / Search time 140.094 Seconds  
(without alignments)  
1052.855 Million cell1 updates/sec

Title: US-09-634-252A-4\_COPY\_532\_586

Perfect score: 332

Sequence: 1 CSLSNGAHCSDGPCNNNTSC.....CDITEYCTGDSGQCPRNLHK 55

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1792395 seqs, 134090451 residues

Total number of hits satisfying chosen parameters: 3504750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US09634222/runat\_21102003\_090327\_3499/app\_query.fasta\_1.462  
-DB=Published Applications\_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=110sum62  
-TRANS=numan40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MTLEN=0  
-MAXLEN=200000000 -USER=US09634252 @CGN 1 1 339 @runat\_21102003\_090327\_3499  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -HSPLOCK=100  
-LONCLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	1668	9	US-09-792-200B-17 Sequence 17, Appl

2	332	100.0	2268	9	US-09-809-790-1	Sequence 1, Appl1
3	332	100.0	2268	10	US-09-809-617-1	Sequence 1, Appl1
4	332	100.0	2268	12	US-10-260-506-1	Sequence 1, Appl1
5	332	100.0	2499	14	US-10-202-675-2	Sequence 2, Appl1
6	332	100.0	4043	9	US-09-824-129-2	Sequence 2, Appl1
7	332	100.0	4043	12	US-10-359-464-2	Sequence 2, Appl1
8	332	100.0	4720	12	US-10-260-506-9	Sequence 9, Appl1
9	325	97.9	494	11	US-09-918-995-22101	Sequence 22101, A
10	285	85.8	2088	9	US-09-809-790-3	Sequence 3, Appl1
11	285	85.8	2088	10	US-09-809-617-3	Sequence 3, Appl1
12	285	85.8	2088	12	US-10-260-506-3	Sequence 3, Appl1
13	285	85.8	2127	12	US-10-260-506-11	Sequence 11, Appl1
14	213	64.2	494	11	US-09-918-995-2610	Sequence 2610, Ap
15	185	55.7	1674	9	US-09-792-200B-15	Sequence 15, Appl1
16	185	55.7	2604	9	US-10-202-675-5	Sequence 7, Appl1
17	168.5	50.8	1638	9	US-09-792-200B-7	Sequence 230, App
18	168.5	50.8	2720	12	US-10-101-510-230	Sequence 3359, Ap
19	168.5	50.8	2740	12	US-09-880-107-3359	Sequence 418, App
20	168.5	50.8	2740	12	US-10-101-510-418	Sequence 1, Appl1
21	168.5	50.8	2968	13	US-10-014-501-1	Sequence 1, Appl1
22	167	50.3	2373	12	US-10-156-028-1	Sequence 203, App
23	167	50.3	2695	12	US-10-137-870-203	Sequence 203, App
24	167	50.3	2695	12	US-10-140-018-203	Sequence 203, App
25	167	50.3	2695	12	US-10-140-021-203	Sequence 203, App
26	167	50.3	2695	12	US-10-140-274-203	Sequence 203, App
27	167	50.3	2695	12	US-10-140-807-203	Sequence 203, App
28	167	50.3	2695	12	US-10-140-471-203	Sequence 203, App
29	167	50.3	2695	12	US-10-140-922-203	Sequence 203, App
30	167	50.3	2695	12	US-10-140-924-203	Sequence 203, App
31	167	50.3	2695	12	US-10-140-926-203	Sequence 203, App
32	167	50.3	2695	12	US-10-141-698-203	Sequence 203, App
33	167	50.3	2695	12	US-10-141-702-203	Sequence 203, App
34	167	50.3	2695	12	US-10-141-704-203	Sequence 203, App
35	167	50.3	2695	12	US-10-142-421-203	Sequence 203, App
36	167	50.3	2695	12	US-10-142-432-203	Sequence 203, App
37	167	50.3	2695	12	US-10-142-767-203	Sequence 203, App
38	167	50.3	2695	12	US-10-143-033-203	Sequence 203, App
39	167	50.3	2695	12	US-10-144-994-203	Sequence 203, App
40	167	50.3	2695	12	US-10-145-628-203	Sequence 203, App
41	167	50.3	2695	12	US-10-145-631-203	Sequence 203, App
42	167	50.3	2695	12	US-10-145-633-203	Sequence 203, App
43	167	50.3	2695	12	US-10-145-746-203	Sequence 203, App
44	167	50.3	2695	12	US-10-145-748-203	Sequence 203, App
45	167	50.3	2695	12	US-10-145-823-203	Sequence 203, App

ALIGNMENTS

RESULT 1  
US-09-792-200B-17  
Sequence 17, Application US/09792200B  
Patent No. US20020042368A1  
GENERAL INFORMATION:  
APPLICANT: Immunex Corporation  
APPLICANT: Fanslow, William C.  
APPLICANT: Poindexter, Kurt  
APPLICANT: Cerretti, Douglas P.  
TITLE OF INVENTION: INTEGRIN ANTAGONISTS  
FILE REFERENCE: 2958-A  
CURRENT APPLICATION NUMBER: US/09/792, 200B  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 60/184,865  
PRIOR FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 1668  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fusion Construct  
NAME/KEY: CDS

```
; LOCATION: (25)..(1647)
; OTHER INFORMATION:
US-09-792-200B-17

Alignment Scores:
Pred. No.: 1,22e-36      Length: 1668
Score: 332.00           Matches: 55
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 9                    Gaps: 0

US-09-634-252A-4_COPY_532_586 (1-55) x US-09-792-200B-17 (1-1668)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThiSerCys 20
DB 172 TGTTCCTCTCCACAGGGGCTCCTGCACGACGGGCTCTGTGAACATATCTCATGT 231
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
DB 232 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATTAATCTGAA 291
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
DB 292 TATTGTACTGGAGACTCTGTGTAGTGCACCAAAATCTTCATTAAG 336

RESULT 2
US-09-809-790-1
; Sequence 1, Application US/09809790
; Patent No. US20020072102A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/809,790
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2090)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2268)
; OTHER INFORMATION: n = A,T,C or G
US-09-809-790-1
```

```
Alignment Scores:
Pred. No.: 1,79e-36      Length: 2268
Score: 332.00           Matches: 55
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 9                    Gaps: 0

US-09-634-252A-4_COPY_532_586 (1-55) x US-09-809-790-1 (1-2268)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThiSerCys 20
DB 1227 TGTTCCTCTCCACAGGGGCTCCTGCACGACGGGCTCTGTGAACATATCTCATGT 1286
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
DB 1287 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATTAATCTGAA 1346
```

```
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
DB 1347 TATTGTACTGGAGACTCTGTGTAGTGCACCAAAATCTTCATTAAG 1391

RESULT 3
US-09-809-617-1
; Sequence 1, Application US/09809617
; Patent No. US20020137178A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/809,617
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: US/09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2090)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2268)
; OTHER INFORMATION: n = A,T,C or G
US-09-809-617-1
```

```
Alignment Scores:
Pred. No.: 1,79e-36      Length: 2268
Score: 332.00           Matches: 55
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 10                   Gaps: 0
```

```
US-09-634-252A-4_COPY_532_586 (1-55) x US-09-809-617-1 (1-2268)

QY 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThiSerCys 20
DB 1227 TGTTCCTCTCCACAGGGGCTCCTGCACGACGGGCTCTGTGAACATATCTCATGT 1286
QY 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
DB 1287 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATTAATCTGAA 1346
QY 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisLys 55
DB 1347 TATTGTACTGGAGACTCTGTGTAGTGCACCAAAATCTTCATTAAG 1391
```

```
RESULT 4
US-10-260-506-1
; Sequence 1, Application US/10260506
; Publication No. US20030153064A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Baidur, Nand
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Taft, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/10/260,506
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410
; PRIOR FILING DATE: 1999-08-05
```



```

; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2090)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2268)
; OTHER INFORMATION: n = A,T,C or G
US-10-260-506-1

Alignment Scores:
Pred. No.: 1,79e-36 Length: 2268
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-10-260-506-1 (1-2268)

Qy 1 CysSerLeuSerArgnGlyAlaHisCysSerAspGlyProCysGlyAsnAsnThrSerCys 20
Db 1227 TGTTCCTCTCCACGGGGCTCACTGCAGCGAGCGGCGCTCTGTAAACAATACCTCATGT 1286

Qy 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db 1287 CTTTTCAGCCACGAGGATGATGATGCCGGGATCTGTGAACGATGATATTACTGAA 1346

Qy 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55
Db 1347 TATTGTACTGAGACTGTGTGCTAGTCCACCAATCTTCAATAG 1391

RESULT 5
US-10-202-675-2
; Sequence 2, Application US/10202675
; Publication No. US20030100091A1
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES
; FILE REFERENCE: 03260,0051-00304
; CURRENT APPLICATION NUMBER: US/10/202,675
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/634,252
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/074,310
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2499
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-202-675-2

Alignment Scores:
Pred. No.: 2,03e-36 Length: 2499
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-10-202-675-2 (1-2499)

Qy 1 CysSerLeuSerArgnGlyAlaHisCysSerAspGlyProCysGlyAsnAsnThrSerCys 20

```

```

Db 1594 TGTTCCTCTCCACGGGGCTCACTGCAGCGAGCGGCGCTCTGTAAACAATACCTCATGT 1653

Qy 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db 1654 CTTTTCAGCCACGAGGATGATGATGCCGGGATCTGTGAACGATGATATTACTGAA 1713

Qy 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55
Db 1714 TATTGTACTGAGACTGTGTGCTAGTCCACCAATCTTCAATAG 1758

RESULT 6
US-09-824-129-2
; Sequence 2, Application US/09824129
; Patent No. US20020001840A1
; GENERAL INFORMATION:
; APPLICANT: Lopez-Otin, Carlos
; APPLICANT: Miguel, Santiago Cal
; APPLICANT: Freije, Jose Maria Perez
; APPLICANT: Garcia, Jose Manuel Lopez
; APPLICANT: Bianchi, Albert Bernard
; APPLICANT: Trail, Pamela
; TITLE OF INVENTION: Methods and Compositions for Modulating
; TITLE OF INVENTION: Integrin-mediated Cell-Cell Interactions
; FILE REFERENCE: D0015-NP
; CURRENT APPLICATION NUMBER: US/09/824,129
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-824-129-2

Alignment Scores:
Pred. No.: 3,72e-36 Length: 4043
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-09-824-129-2 (1-4043)

Qy 1 CysSerLeuSerArgnGlyAlaHisCysSerAspGlyProCysGlyAsnAsnThrSerCys 20
Db 2610 TGTTCCTCTCCACGGGGCTCACTGCAGCGAGCGGCGCTCTGTAAACAATACCTCATGT 2669

Qy 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrGlu 40
Db 2670 CTTTTCAGCCACGAGGATGATGATGCCGGGATCTGTGAACGATGATATTACTGAA 2729

Qy 41 TyrCysThrGlyAspSerGlyGlnCysProProAsnLeuHisIlys 55
Db 2730 TATTGTACTGAGACTGTGTGCTAGTCCACCAATCTTCAATAG 2774

RESULT 7
US-10-359-464-2
; Sequence 2, Application US/10359464
; Publication No. US20030143692A1
; GENERAL INFORMATION:
; APPLICANT: Lopez-Otin, Carlos
; APPLICANT: Miguel, Santiago Cal
; APPLICANT: Freije, Jose Maria Perez
; APPLICANT: Garcia, Jose Manuel Lopez
; APPLICANT: Bianchi, Albert Bernard
; APPLICANT: Trail, Pamela
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING INTEGRIN-MEDIATED
; TITLE OF INVENTION: CELL-CELL INTERACTIONS
; FILE REFERENCE: D0015 DIV
; CURRENT APPLICATION NUMBER: US/10/359,464

```

```
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 09/824,129
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentn version 3.2
; SEQ ID NO 2
; LENGTH: 4043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-359-464-2

Alignment Scores:
Pred. No.: 3,72e-36 Length: 4043
Score: 332.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-10-359-464-2 (1-4043)

QY 1 CysSerLeuSerAenGlyAlaHisCySerAspGlyProCyCyAsnAsnThrSerCys 20
Db 2610 TGTTCCTCTCCCAACGGGGCTCACTGCAGCGACGGGCTCTGTAAACAATACCTCATGT 2669

QY 21 LeuPheGlnProAArgGlyTYrGluCyArgAspAlaValaAngluCyAspIleThrGlu 40
Db 2670 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATATTACTGAA 2729

QY 41 TYrCyThrGlyAspSerGlyGlnCySerProAsnLeuHisLys 55
Db 2730 TATTGACTGGAGACTGTGTGTCAGTCCCAACCAATCTTCATAAG 2774

RESULT 8
US-10-260-506-9
; Sequence 9, Application US/10260506
; Publication No. US20030153064A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Baindur, Nand
; APPLICANT: Desher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Tafel, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/10/260,506
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 4720
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2129)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4720)
; OTHER INFORMATION: n = A,T,C or G
US-10-260-506-9

Alignment Scores:
Pred. No.: 4,53e-36 Length: 4720
Score: 332.00 Matches: 55
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-10-260-506-9 (1-4720)

QY 1 CysSerLeuSerAenGlyAlaHisCySerAspGlyProCyCyAsnAsnThrSerCys 20
Db 1227 TGTTCCTCTCCCAACGGGGCTCACTGCAGCGACGGGCTCTGTAAACAATACCTCATGT 1286

QY 21 LeuPheGlnProAArgGlyTYrGluCyArgAspAlaValaAngluCyAspIleThrGlu 40
Db 1287 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATATTACTGAA 1346

QY 41 TYrCyThrGlyAspSerGlyGlnCySerProAsnLeuHisLys 55
Db 1347 TATTGACTGGAGACTGTGTGTCAGTCCCAACCAATCTTCATAAG 1391

RESULT 9
US-09-918-995-22101
; Sequence 22101, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22101
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(494)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22101

Alignment Scores:
Pred. No.: 2,49e-36 Length: 494
Score: 325.00 Matches: 54
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 1
Query Match: 97.89% Indels: 0
DB: 11 Gaps: 0

US-09-634-252a-4_COPY_532_586 (1-55) x US-09-918-995-22101 (1-494)

QY 1 CysSerLeuSerAenGlyAlaHisCySerAspGlyProCyCyAsnAsnThrSerCys 20
Db 234 TGTTCCTCTCCCAACGGGGCTCACTGCAGCGACGGGCTCTGTAAACAATACCTCATGT 293

QY 21 LeuPheGlnProAArgGlyTYrGluCyArgAspAlaValaAngluCyAspIleThrGlu 40
Db 294 CTTTTCAGCCACGAGGGATGAAATGCCGGATGCTGTGAACGAGTGTGATATTACTGAA 353

QY 41 TYrCyThrGlyAspSerGlyGlnCySerProAsnLeuHisLys 55
Db 354 TATTGACTGGAGACTGTGTGTCAGTCCCAACCAATCTTCATAAG 398

RESULT 10
US-09-809-790-3
; Sequence 3, Application US/09809790
; Patent No. US20020072102A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Baindur, Nand
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:10:45 ; Search time 110.692 Seconds  
(without alignments)  
83.207 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586

Perfect score: 332  
Sequence: 1 CSLSNGAHCSDBGPCNNNTSC.....CDITEYCTGDSGCCPPNLHK 55

Scoring table:  
BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgnt\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgnt\_6/ptodata/1/pubppa/PTCT\_NEW\_PUB.pep:\*  
3: /cgnt\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgnt\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgnt\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgnt\_6/ptodata/1/pubppa/CTUS\_PUBCOMB.pep:\*  
7: /cgnt\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgnt\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgnt\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgnt\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgnt\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgnt\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
13: /cgnt\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgnt\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgnt\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep:\*  
16: /cgnt\_6/ptodata/1/pubppa/US10C\_NEW\_PUB.pep:\*  
17: /cgnt\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
18: /cgnt\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332	100.0	540	US-09-792-2008-18	Sequence 18, Appl
2	332	100.0	696	US-09-809-790-2	Sequence 2, Appl1
3	332	100.0	696	US-09-809-617-2	Sequence 2, Appl1
4	332	100.0	696	US-10-260-506-2	Sequence 2, Appl1
5	332	100.0	709	US-10-260-506-10	Sequence 10, Appl
6	332	100.0	832	US-09-824-129-3	Sequence 3, Appl1
7	332	100.0	832	US-10-359-464-3	Sequence 3, Appl1
8	332	100.0	832	US-10-202-675-4	Sequence 4, Appl1
9	185	55.7	542	US-09-792-2008-16	Sequence 16, Appl
10	185	55.7	867	US-10-202-675-6	Sequence 6, Appl1
11	172	51.8	76	US-09-877-843-88	Sequence 88, Appl
12	168.5	50.8	522	US-09-792-2008-8	Sequence 8, Appl1
13	168.5	50.8	814	US-10-014-501-4	Sequence 4, Appl1
14	168.5	50.8	855	US-10-014-501-2	Sequence 2, Appl1
15	167	50.3	790	US-10-137-870-204	Sequence 204, App

16	167	50.3	790	12	US-10-140-018-204	Sequence 204, App
17	167	50.3	790	12	US-10-140-021-204	Sequence 204, App
18	167	50.3	790	12	US-10-140-274-204	Sequence 204, App
19	167	50.3	790	12	US-10-140-471-204	Sequence 204, App
20	167	50.3	790	12	US-10-140-807-204	Sequence 204, App
21	167	50.3	790	12	US-10-140-922-204	Sequence 204, App
22	167	50.3	790	12	US-10-140-924-204	Sequence 204, App
23	167	50.3	790	12	US-10-140-926-204	Sequence 204, App
24	167	50.3	790	12	US-10-141-698-204	Sequence 204, App
25	167	50.3	790	12	US-10-141-702-204	Sequence 204, App
26	167	50.3	790	12	US-10-141-704-204	Sequence 204, App
27	167	50.3	790	12	US-10-142-421-204	Sequence 204, App
28	167	50.3	790	12	US-10-142-432-204	Sequence 204, App
29	167	50.3	790	12	US-10-142-767-204	Sequence 204, App
30	167	50.3	790	12	US-10-143-033-204	Sequence 204, App
31	167	50.3	790	12	US-10-144-994-204	Sequence 204, App
32	167	50.3	790	12	US-10-145-628-204	Sequence 204, App
33	167	50.3	790	12	US-10-145-631-204	Sequence 204, App
34	167	50.3	790	12	US-10-145-633-204	Sequence 204, App
35	167	50.3	790	12	US-10-145-746-204	Sequence 204, App
36	167	50.3	790	12	US-10-145-748-204	Sequence 204, App
37	167	50.3	790	12	US-10-145-823-204	Sequence 204, App
38	167	50.3	790	12	US-10-145-826-204	Sequence 204, App
39	167	50.3	790	12	US-10-145-876-204	Sequence 204, App
40	167	50.3	790	12	US-10-145-876-204	Sequence 204, App
41	167	50.3	790	12	US-10-145-959-204	Sequence 204, App
42	167	50.3	790	12	US-10-146-724-204	Sequence 204, App
43	167	50.3	790	12	US-10-146-725-204	Sequence 204, App
44	167	50.3	790	12	US-10-146-795-204	Sequence 204, App
45	167	50.3	790	12	US-10-147-495-204	Sequence 204, App

## ALIGNMENTS

RESULT 1  
US-09-792-2008-18 Application US/09792200B  
Sequence 18, Appl1  
Patent No. US20020042368A1  
GENERAL INFORMATION:  
APPLICANT: Immunex Corporation  
APPLICANT: Fanslow, William C.  
APPLICANT: Poindexter, Kurt  
APPLICANT: Corelli, Douglas P.  
APPLICANT: Black, Roy A.  
TITLE OF INVENTION: INTEGRIN ANTAGONISTS  
FILE REFERENCE: 2958-A  
CURRENT APPLICATION NUMBER: US/09/792,200B  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 66/164,865  
PRIOR FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 18  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fusion Construct  
US-09-792-2008-18  
Query Match 100.0%; Score 332; DB 9; Length 540;  
Best Local Similarity 100.0%; Pred. No. 1.5e-27;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CSLSNGAHCSDBGPCNNNTSCFQPGRYECRDVAVECDITEYCTGDSGCCPPNLHK 55  
DB 50 CSLSNGAHCSDBGPCNNNTSCFQPGRYECRDVAVECDITEYCTGDSGCCPPNLHK 104  
RESULT 2  
US-09-809-790-2  
Sequence 2, Appl1  
Application US/09809790

```
Patent No. US20020072102A1
GENERAL INFORMATION:
APPLICANT: Shepard, Paul O.
APPLICANT: Baidur, Nand
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: DISINTEGRIN HOMOLOG
FILE REFERENCE: 98-29
CURRENT APPLICATION NUMBER: US/09/809,790
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 09/351,414
PRIOR FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 696
TYPE: PRT
ORGANISM: Homo sapiens
US-09-809-790-2
```

```
Query Match          100.0%; Score 332; DB 9; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
|||||
DB 409 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 463
```

```
RESULT 3
US-09-809-617-2
Sequence 2, Application US/09809617
Patent No. US2002037178A1
GENERAL INFORMATION:
APPLICANT: Shepard, Paul O.
APPLICANT: Baidur, Nand
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: DISINTEGRIN HOMOLOG
FILE REFERENCE: 98-29
CURRENT APPLICATION NUMBER: US/09/809,617
CURRENT FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: US/09/351,414
PRIOR FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 696
TYPE: PRT
ORGANISM: Homo sapiens
US-09-809-617-2
```

```
Query Match          100.0%; Score 332; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
|||||
DB 409 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 463
```

```
RESULT 4
US-10-260-506-2
Sequence 2, Application US/10260506
Publication No. US20030153064A1
```

```
GENERAL INFORMATION:
APPLICANT: Shepard, Paul O.
APPLICANT: Baidur, Nand
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
APPLICANT: Taft, David W.
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
FILE REFERENCE: 98-29C1
```

```
CURRENT APPLICATION NUMBER: US/10/260,506
CURRENT FILING DATE: 2002-09-27
PRIOR APPLICATION NUMBER: 60/092,371
PRIOR FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 60/147, 410
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 09/351,414
PRIOR FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 696
TYPE: PRT
ORGANISM: Homo sapiens
US-10-260-506-2
```

```
Query Match          100.0%; Score 332; DB 12; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
|||||
DB 409 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 463
```

```
RESULT 5
US-10-260-506-10
Sequence 10, Application US/10260506
Publication No. US20030153064A1
GENERAL INFORMATION:
APPLICANT: Shepard, Paul O.
APPLICANT: Baidur, Nand
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
APPLICANT: Taft, David W.
TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
FILE REFERENCE: 98-29C1
CURRENT APPLICATION NUMBER: US/10/260,506
CURRENT FILING DATE: 2002-09-27
PRIOR APPLICATION NUMBER: 60/092,371
PRIOR FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 60/147, 410
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 09/351,414
PRIOR FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
US-10-260-506-10
```

```
Query Match          100.0%; Score 332; DB 12; Length 709;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
|||||
DB 409 CSLSNAGHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 463
```

```
RESULT 6
US-09-824-129-3
```

```
Sequence 3, Application US/09824129
Patent No. US2002001840A1
GENERAL INFORMATION:
APPLICANT: Lopez-Otin, Carlos
APPLICANT: Miguel, Santiago Cal
APPLICANT: Freije, Jose Maria Perez
APPLICANT: Garcia, Jose Manuel Lopez
APPLICANT: Bianchi, Albert Bernard
APPLICANT: Trall, Pamela
```

```
; TITLE OF INVENTION: Methods and Compositions for Modulating
; FILE REFERENCE: D0015-NP
; CURRENT APPLICATION NUMBER: US/09/824,129
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-824-129-3

Query Match      100.0%; Score 332; DB 9; Length 832;
Best Local Similarity 100.0%; Pred. No. 2,3e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
Db      532 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 586
```

```
RESULT 7
US-10-359-464-3
; Sequence 3, Application US/10359464
; Publication No. US20030143692A1
; GENERAL INFORMATION:
; APPLICANT: Lopez-Otin, Carlos
; APPLICANT: Migue, Santiago Cal
; APPLICANT: Freije, Jose Maria Perez
; APPLICANT: Garcia, Jose Manuel Lopez
; APPLICANT: Bianchi, Albert Bernard
; APPLICANT: Trail, Pamela
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING INTEGRIN-MEDIATED
; FILE REFERENCE: D0015 DIV
; CURRENT APPLICATION NUMBER: US/10/359,464
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 09/824,129
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-359-464-3

Query Match      100.0%; Score 332; DB 12; Length 332;
Best Local Similarity 100.0%; Pred. No. 2,3e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
Db      532 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 586
```

```
RESULT 8
US-10-202-675-4
; Sequence 4, Application US/10202675
; Publication No. US20030100091A1
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES
; FILE REFERENCE: 03260,0051-00304
; CURRENT APPLICATION NUMBER: US/10/202,675
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/634,252
; PRIOR FILING DATE: 2000-08-07
```

```
; PRIOR APPLICATION NUMBER: 60/074,310
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-202-675-4
```

```
Query Match      100.0%; Score 332; DB 15; Length 832;
Best Local Similarity 100.0%; Pred. No. 2,3e-27;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
Db      532 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 586
```

```
RESULT 9
US-09-792-200B-16
; Sequence 16, Application US/09792200B
; Patent No. US20020042368A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Fanslow, William C.
; APPLICANT: Poindexter, Kurt
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Black, Roy A.
; TITLE OF INVENTION: INTEGRIN ANTAGONISTS
; FILE REFERENCE: 2958-A
; CURRENT APPLICATION NUMBER: US/09/792,200B
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/184,865
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion Construct
US-09-792-200B-16
```

```
Query Match      55.7%; Score 185; DB 9; Length 542;
Best Local Similarity 56.4%; Pred. No. 6,6e-12;
Matches 31; Conservative 8; Mismatches 14; Indels 2; Gaps 1;
```

```
Oy      1 CSLSNGAHCSGDPCCNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55
Db      53 CILTQDSQSDGLCK--KCFQPMGTVCREAVNDCDIRETCGSSQCAPNLHK 105
```

```
RESULT 10
US-10-202-675-6
; Sequence 6, Application US/10202675
; Publication No. US20030100091A1
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES
; FILE REFERENCE: 03260,0051-00304
; CURRENT APPLICATION NUMBER: US/10/202,675
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/09/634,252
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/074,310
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 867
; TYPE: PRT
```







```

; Sequence 2, Application US/09809617
; Patent No. US20020137178A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baird, Nand O.
; APPLICANT: Delsher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/809,617
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: US/09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-617-2

```

```

Query Match          100.0%; Score 630; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.9e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 60
    |||||
DB 373 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 432
    |||||
QY 61 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 104
    |||||
DB 433 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 476
    |||||

```

```

RESULT 3
US-10-260-506-2
; Sequence 2, Application US/10260506
; Patent No. US20030153064A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baird, Nand O.
; APPLICANT: Delsher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Taft, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/10/260,506
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-260-506-2

```

```

Query Match          100.0%; Score 630; DB 12; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.9e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 60
    |||||
DB 373 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 432
    |||||
QY 61 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 104
    |||||
DB 433 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 476
    |||||

```

```

RESULT 4
US-10-260-506-10
; Sequence 10, Application US/10260506
; Patent No. US20030153064A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baird, Nand O.
; APPLICANT: Delsher, Theresa A.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Taft, David W.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOGUE, MAHBP
; FILE REFERENCE: 98-29C1
; CURRENT APPLICATION NUMBER: US/10/260,506
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/092,371
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/147, 410
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 09/351,414
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-260-506-10

```

```

Query Match          100.0%; Score 630; DB 12; Length 709;
Best Local Similarity 100.0%; Pred. No. 2e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 60
    |||||
DB 373 PTKLFPEPTCGNGYVAGEECDCGFHVECYGLCKCKKSLSGAHGSDGFCNNNTSCLFOP 432
    |||||
QY 61 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 104
    |||||
DB 433 RGYECRDVAIVEDITETCTGDSGGCCPPLHKGODGYACNONGRC 476
    |||||

```

```

RESULT 5
US-09-824-129-3
; Sequence 3, Application US/09824129
; Patent No. US20020001840A1
; GENERAL INFORMATION:
; APPLICANT: Lopez-Otin, Carlos
; APPLICANT: Miguel, Santiago Cal
; APPLICANT: Freije, Jose Maria Perez
; APPLICANT: Garcia, Jose Manuel Lopez
; APPLICANT: Bianchi, Albert Bernard
; APPLICANT: Trall, Pamela
; TITLE OF INVENTION: Methods and Compositions for Modulating
; FILE REFERENCE: D0015-NP
; CURRENT APPLICATION NUMBER: US/09/824,129
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/194,164
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-824-129-3

```

```

Query Match          100.0%; Score 630; DB 9; Length 832;
Best Local Similarity 100.0%; Pred. No. 2.3e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 PTKLFEEPTGCGNGVYGEAGECDGCFHNECYGLCCCKKLSLNGAH(S)GCPCCNNTSCLFOP 60

Db 496 PTKLFEEPTGCGNGVYGEAGECDGCFHNECYGLCCCKKLSLNGAH(S)GCPCCNNTSCLFOP 555

QY 61 RGAYCRDAVNECDITTEYCTGDSGGCPNLHKQDSYAQNONGAGC 104

Db 556 RGAYCRDAVNECDITTEYCTGDSGGCPNLHKQDSYAQNONGAGC 599

RESULT 6

```

US-10-359-464-3
: Sequence 3, Application US/10359464
: Publication No. US20030143692A1
: GENERAL INFORMATION:
: APPLICANT: Lopez-Otin, Carlos
: APPLICANT: Miguel, Santiago Cal
: APPLICANT: Freije, Jose Maria, Perez
: APPLICANT: Garcia, Jose Manuel Lopez
: APPLICANT: Bianchi, Albert Bernard
: APPLICANT: Trill, Pamela
: TITLE OF INVENTION: METHODS AND COMPOSITIONS
: TITLE OF INVENTION: CELL-CELL INTERACTIONS
: FILE REFERENCE: D0015 DIV
: CURRENT APPLICATION NUMBER: US/10/359,464
: CURRENT FILING DATE: 2003-02-05
: PRIOR APPLICATION NUMBER: 09/824,129
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 60/194,164
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 3
: LENGTH: 832
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-359-464-3

```

Query Match	100.0%	Score 630	DB 12	Length 332
Best Local Similarity	100.0%	Pred. No. 2.3e-48		
Matches	104	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	1	PTKLFEPIECNGGYEAGEECDCGFHVEVCYGLCKKCSISNGAHCSDGZCCNNTSCLFOP	60	
Db	496	PTKLFEPIECNGGYEAGEECDCGFHVEVCYGLCKKCSISNGAHCSDGZCCNNTSCLFOP	555	
QY	61	RGYECRDVNECDITVEYCTGDSGQCPPLIHKDYGACNONQRC	104	
Db	556	RGYECRDVNECDITVEYCTGDSGQCPPLIHKDYGACNONQRC	599	

## RESULT 7

US-10-202-675-4  
 : Sequence 4, Application US/10202675  
 : Publication No. US20030100091A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Cerrett, Douglas P.  
 : TITLE OF INVENTION: SYPH3-13 AND SYPH3-17 DNA AND POLYPEPTIDES  
 : FILE REFERENCE: 03260, 0051-00304  
 : CURRENT APPLICATION NUMBER: US/10/202,675  
 : CURRENT FILING DATE: 2002-07-23  
 : PRIOR APPLICATION NUMBER: US/09/634,252  
 : PRIOR FILING DATE: 2000-08-07  
 : PRIOR APPLICATION NUMBER: 60/074,310  
 : PRIOR FILING DATE: 1998-02-11  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 4  
 : LENGTH: 832  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 : US-10-202-675-4

Query Match 100.0%; Score 630; DB 15; Length 632;

Best Local Similarity 100.0%; Pred. No. 2.3e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	PTKLFPEPCNCNGYVEAEBCDCGFHVEYGLCKCKCSLSNNAHCSDDGCCNNNTSCLFOP	60
Db	496	PTKLFPEPCNCNGYVEAEBCDCGFHVEYGLCKCKCSLSNNAHCSDDGCCNNNTSCLFOP	555
QY	61	RGYECRDVNECDITLEYCTGDSGGCPPLNLRKDDGYACNNNOGR	104
Db	556	RGYECRDVNECDITLEYCTGDSGGCPPLNLRKDDGYACNNNOGR	599

RESULT 8

```

: Sequence 18, Application US/09792200B
: Patent No. US20020042368A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Immunex Corporation
: APPLICANT: Fanslow, William C.
: APPLICANT: Poindexter, Kurt
: APPLICANT: Cerretti, Douglas P.
: APPLICANT: Black, Roy A.
: TITLE OF INVENTION: INTEGRIN ANTAGONISTS
: FILE REFERENCE: 2958-A
: CURRENT APPLICATION NUMBER: US/09/792,200B
: CURRENT FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: US 60/184,865
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: SeqId version 3.1
:
: SEQ ID NO 18
:
: LENGTH: 540
:
: TYPE: PRT
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Fusion Construct
:
: US-09-792-200B-18

```

Query Match	93.0%	Score 586	DB 9	Length 540
Best Local Similarity	99.0%	Pred. No. 1.3e+4		
Matches	96	Conservative	0	Mismatches 1
				Indels 0
				Gaps 0
Qy	8	TECGNGYAEAEBCDCGFHVECTYGLCKCKKSLNSGAHCSDDPCCNNTSCLFOPRAGECRD	67	
db	21	TSCGNGYAEAEBCDCGFHVECTYGLCKCKKSLNSGAHCSDDPCCNNTSCLFOPRAGECRD	80	
Qy	68	AVNECDITREYCTGDSGCCPPNMLHKODGYACACNQNQSRC	104	
Db	81	AVNECDITREYCTGDSGCCPPNMLHKODGYACACNQNQSRC	117	

## RESULT 9

```

US-10-202-675-6
: Sequence 6, Application US/10202675
: Publication No. US20030100091A1
: GENERAL INFORMATION:
: APPLICANT: Cerretti, Douglas P.
: TITLE OF INVENTION: SVPH3-13 AND SVPH3-17 DNA AND POLYPEPTIDES
: FILE REFERENCE: 03260, 0051-00304
: CURRENT APPLICATION NUMBER: US/10/202,675
: CURRENT FILING DATE: 2002-07-23
: PRIOR APPLICATION NUMBER: US/09/634,252
: PRIOR FILING DATE: 2000-08-07
: PRIOR APPLICATION NUMBER: 60/074,310
: PRIOR FILING DATE: 1998-02-11
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 6
: LENGTH: 867
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-202-675-6

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:20:51 ; Search time 32.8616 Seconds  
(without alignments)  
738.736 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586  
Perfect score: 332  
Sequence: 1 CSLSNGAHGSDGPCNNNTSC.....CDITEXCTGDSGQCPNHLK 55

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Fgapext 0.5  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139955

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+.p2n.model -DEV=x1h  
-Q/cgn2\_1/USFTO.spool/US09634452/rnat\_21102003\_090325\_3401/app\_query.fasta\_1.462  
-DB=Issued\_Patents\_NA -OPMT=faetext -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=10  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09634452 @CCN 1 1 76 @rnat\_21102003\_090325\_3401 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCL=100 -LONGCLOS  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	332	100.0	2268	3 US-09-351-414-1	Sequence 1, Appl1
2	285	85.8	2088	3 US-09-351-414-3	Sequence 5, Appl1
3	213	64.2	1464	1 US-08-243-542-5	Sequence 5, Appl1
4	213	64.2	1464	1 US-08-477-407-5	Sequence 5, Appl1
5	213	64.2	1464	1 US-08-484-355-5	Sequence 5, Appl1
6	213	64.2	2913	1 US-08-243-542-7	Sequence 7, Appl1
7	213	64.2	2913	1 US-08-477-407-7	Sequence 7, Appl1
8	213	64.2	2913	1 US-08-484-355-7	Sequence 7, Appl1
9	213	64.2	2923	1 US-08-243-542-6	Sequence 6, Appl1
10	213	64.2	2923	1 US-08-477-407-6	Sequence 6, Appl1
11	213	64.2	2923	1 US-08-484-355-6	Sequence 6, Appl1
12	213	64.2	3183	1 US-08-243-542-8	Sequence 8, Appl1

13	213	64.2	3183	1 US-08-477-407-8	Sequence 8, Appl1
14	213	64.2	3183	1 US-08-484-355-8	Sequence 8, Appl1
15	168.5	50.8	2968	3 US-09-813-819-1	Sequence 1, Appl1
16	168.5	50.8	2968	4 US-09-920-048-1	Sequence 1, Appl1
17	160	48.2	1507	4 US-09-280-116-122	Sequence 122, App
18	160	48.2	3431	4 US-09-632-098-1	Sequence 1, Appl1
19	160	48.2	3468	4 US-09-632-098-3	Sequence 3, Appl1
20	153	46.1	1463	5 PCT-US93-05640-10	Sequence 10, Appl1
21	152	45.8	2373	1 US-08-264-101-1	Sequence 1, Appl1
22	152	45.8	2373	2 US-08-765-243-1	Sequence 1, Appl1
23	152	45.8	2373	5 PCT-US95-07295-1	Sequence 1, Appl1
24	152	45.8	2373	5 PCT-US95-07295-7	Sequence 7, Appl1
25	152	45.8	2650	5 PCT-US95-07295-7	Sequence 7, Appl1
26	149	44.9	485	4 US-09-280-116-39	Sequence 39, Appl1
27	147	44.3	1851	4 US-09-608-790-2	Sequence 2, Appl1
28	145	43.7	2169	4 US-09-617-145-1	Sequence 1, Appl1
29	144.5	43.5	3994	4 US-09-738-946-7	Sequence 7, Appl1
30	144	43.4	1768	1 US-08-264-101-3	Sequence 3, Appl1
31	144	43.4	1768	2 US-08-765-243-3	Sequence 3, Appl1
32	144	43.4	1768	5 PCT-US95-07295-3	Sequence 3, Appl1
33	144	43.4	2553	2 US-08-765-243-5	Sequence 5, Appl1
34	144	43.4	2553	5 PCT-US95-07295-5	Sequence 5, Appl1
35	136	41.0	2050	4 US-09-026-001A-5	Sequence 5, Appl1
36	135	40.7	2406	4 US-09-632-098-5	Sequence 5, Appl1
37	135	40.7	2439	4 US-09-632-098-6	Sequence 6, Appl1
38	132.5	39.9	2097	2 US-08-655-345-3	Sequence 3, Appl1
39	132.5	39.9	2097	4 US-09-183-275-3	Sequence 3, Appl1
40	132.5	39.9	2097	4 US-09-727-169-3	Sequence 3, Appl1
41	132.5	39.9	2087	4 US-09-579-766A-3	Sequence 3, Appl1
42	132.5	39.9	2097	4 US-09-726-968-3	Sequence 3, Appl1
43	132.5	39.9	2097	5 PCT-US96-08407-3	Sequence 3, Appl1
44	132.5	39.9	2475	2 US-08-655-345-1	Sequence 1, Appl1
45	132.5	39.9	2475	3 US-09-183-275-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-09-351-414-1  
; Sequence 1, Application US/09351414  
; Patent No. 6265199  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Baird, Paul  
; APPLICANT: Deisher, Theresa A.  
; APPLICANT: Bishop, Paul D.  
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
; FILE REFERENCE: 98-29  
; CURRENT APPLICATION NUMBER: US/09/351,414  
; CURRENT FILING DATE: 1999-07-09  
; NUMBER OF SEQ. ID NOS: 13  
; SOFTWARE: FASTA/Seq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2268  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)..(2090)  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(2268)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-351-414-1  
Alignment Scores:  
Pred. No.: 2,33e-33  
Score: 332.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 3  
Length: 2268  
Matches: 55  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0  
US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-09-351-414-1 (1-2268)

Qy 1 CysSerLeuSerAanglyAlaHisCysSerAaspGlyProCysCysAsnAsnThSerCys 20  
Db 1227 TGTTCCTCTCCACAGCGGGCTACTGACGCGCGCCCTCTGTAAACAATACCTCATGT 1286  
Qy 21 LeuPheGlnProAArgGlyTyTGluCysArgAspAlaValaAngGluCysAsp11eThrglu 40  
Db 1287 CTTTTCAGCCACGAGGGGTATGAATGCCGGAGTGTGAAAGAGTGTGATATTACTGAA 1346  
Qy 41 TyrCysThrglyAaspSerGlyGlnCysProProAsnLeuHisLys 55  
Db 1347 TATTGTACTGAGACTCTGTCTGACAGTCCGCCCAATCTTCAATAG 1391  
RESULT 2  
US-09-351-414-3  
Sequence 3, Application US/09351414  
Patent No. 6265199  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Baidur, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/351,414  
CURRENT FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 2088  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: zdmnt amino acid degenerate sequence  
NAME/KEY: variation  
LOCATION: (1)...(2088)  
FEATURE:  
OTHER INFORMATION: n is any nucleotide  
NAME/KEY: misc feature  
LOCATION: (1)...(2088)  
OTHER INFORMATION: n = A,T,C or G  
US-09-351-414-3  
Alignment Scores:  
Pred. No.: 2,31e-27 Length: 2088  
Score: 285.00 Matches: 45  
Percent Similarity: 81.82% Conservative: 0  
Best Local Similarity: 81.82% Mismatches: 10  
Query Match: 85.84% Indels: 0  
DB: Gaps: 0  
US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-09-351-414-3 (1-2088)  
Qy 1 CysSerLeuSerAanglyAlaHisCysSerAaspGlyProCysCysAsnAsnThSerCys 20  
Db 1225 TGYMSNTYTMGNAAAYGGGNCACATGYWNGAYGGGNCNTGTGTAAAYAAACNMNSNTGY 1284  
Qy 21 LeuPheGlnProAArgGlyTyTGluCysArgAspAlaValaAngGluCysAsp11eThrglu 40  
Db 1285 YTNNTTCARCCNMGNGNTAAGARTGYNGNGAYGGGNTAAAYGARTGTGAAYATTAACNGAR 1344  
Qy 41 TyrCysThrglyAaspSerGlyGlnCysProProAsnLeuHisLys 55  
Db 1345 TATGTACNGGNGAYWNGNCARCTGCCNCAATYTNCAIYAR 1389  
RESULT 3  
US-08-243-542-5  
Sequence 5, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-243-542-5  
Alignment Scores:  
Pred. No.: 2.83e-18 Length: 1464  
Score: 213.00 Matches: 36  
Percent Similarity: 76.36% Conservative: 6  
Best Local Similarity: 65.45% Mismatches: 11  
Query Match: 64.16% Indels: 2  
DB: Gaps: 1  
US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-08-243-542-5 (1-1464)  
Qy 1 CysSerLeuSerAanglyAlaHisCysSerAaspGlyProCysCysAsnAsnThSerCys 20  
Db 1111 TGCACCTGACTCAGCAGCCCATGTGACGACGCGCCTGTCTCCGCC-----TGC 1164  
Qy 21 LeuPheGlnProAArgGlyTyTGluCysArgAspAlaValaAngGluCysAsp11eThrglu 40  
Db 1165 AAGTACGACACCGGGGTGTCTCTGCGAGAGCCGTGAACAGTGCACATGCGGAG 1224  
Qy 41 TyrCysThrglyAaspSerGlyGlnCysProProAsnLeuHisLys 55

Db 1225 ACTGCACCGGGAGCTAGCCAGTCCCGCCTAAGCTGACAA3 1269

## RESULT 4

US-08-477-407-5

Sequence 5, Application US/08477407  
Patent No. 5631351

## GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994

## ATTORNEY/AGENT INFORMATION:

NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465

## INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: ORGANISM: Homo sapiens  
IMMEDIATE SOURCE: LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464

US-08-477-407-5

## Alignment Scores:

Pred. No.: 2,83e-18 Length: 1464  
Score: 213.00 Matches: 36  
Percent Similarity: 76.36% Conservative: 6  
Best Local Similarity: 65.45% Mismatches: 11  
Query Match: 64.16% Indels: 2  
DB: 1 Gaps: 1

US-09-634-252a-4\_copy\_532\_586 (1-55) x US-08-477-407-5 (1-1464)

Qy 1 CysSerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCys 20

Db 1111 TGCACCTGACTCAGCAGCGCATGTGTCAGCGAGCGGTCTGCTGTCCGCC-----TGC 1164

Qy 21 LeuPheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAsp11eTh-Glu 40

Db 1165 AAGTACGAACACGAGCGGTGTCTCCCGCAGAGCGCTGAACGAGTCGACATCGCGAG 1224

Qy 41 Ty-CysThrGlyAspSerGlyInCysProProAsnLeuHisLys 55

Db 1225 ACTGCACCGGGAGCTAGCCAGTCCCGCCTAAGCTGACAA3 1269

## RESULT 5

US-08-484-355-5

Sequence 5, Application US/08484355  
Patent No. 5705341

## GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994

## ATTORNEY/AGENT INFORMATION:

NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465

## INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: ORGANISM: Homo sapiens  
IMMEDIATE SOURCE: LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464







RESULT 9  
US-08-243-542-6  
Sequence 6, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
NUMBER OF INVENTIONS: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2923 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: 5' UTR  
LOCATION: 1..27  
NAME/KEY: 3' UTR  
LOCATION: 1600..2923  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..1599  
US-08-243-542-6  
Alignment Scores:  
Pred. No.: 6, 9e-18 Length: 2923  
Score: 213.00 Matches: 36  
Percent Similarity: 76.36% Conservative: 6  
Best Local Similarity: 65.45% Mismatches: 11  
Query Match: 64.16% Indels: 2  
DB: 1 Gaps: 1

US-09-634-252A-4\_COPY\_532\_586 (1-55) x US-08-243-542-6 (1-2923)  
QY 1 CysSerLeuSerArganglyAlaHisCysSerSpGlyProCysCysAsnAsnThrSerCys 20  
Db 1159 TGCACCTGACTCAGCAGCGCATGTGCAGCGCGCTGTCTGTCCGCCG-----TGC 1212  
QY 21 LeuPheGlnProAArgGlyTyrGluCysArgAspAlaValaAngluCysAspIleThrGlu 40  
Db 1213 AAGTACGAACACCGGGGTGTCTCCTGCCGAGAGCGCCGTGAACGAGTGGACATCGCGAG 1272  
QY 41 TyrCysThrGlyAaspSerGlyGlnCysProProAsnLeuHisLys 55  
Db 1273 ACCTGACCGGGGACTGACGACAGTGCCTAAGCTGACAG 1317  
RESULT 10  
US-08-477-407-6  
Sequence 6, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
NUMBER OF INVENTIONS: ENCODING THE SAME  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2923 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:

NAME/KEY: 5' UTR  
 LOCATION: 1..27  
 FEATURE:  
 NAME/KEY: 3' UTR  
 LOCATION: 1600..2923  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 28..1599  
 US-08-477-407-6

Alignment Scores:

Pred. No.: 6.9e-18 Length: 2923  
 Score: 213.00 Matches: 36  
 Percent Similarity: 76.36% Conservative: 6  
 Best Local Similarity: 65.45% Mismatches: 11  
 Query Match: 64.16% Indels: 2  
 DB: 1 Gaps: 1

US-09-634-252a-4\_COPY\_532\_586 (1-55) x US-08-477-407-6 (1-2923)

QY 1 CysSerLeuSerAenGlyAlaHisCysSerAenGlyProCysCysAenAenThrSerCys 20  
 Db 1159 TGCACCTGACTCAGACGCCATGTGACGACGCGGCTCTGCTGCGCCG-----TGC 1212  
 QY 21 LeuPheGlnProAArgGlyTyrgluCysAArgAAlaValAenGluCysAspIleThrGlu 40  
 Db 1213 AAGTACGAACCAACGCGGCTGTCTCTGCGAGAGCCGCGTGAACGAGTGGACATCGCGAG 1272  
 QY 41 TyrCysThrGlyAenPheSerGlyGlnCysProProAenLeuHisIle 55  
 Db 1273 ACCTGACCGCGGACTCTAGCCAGTGCCTAACCTGACCAAG 1317

Search completed: October 21, 2003, 12:03:47  
 Job time : 37.8616 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:06:05 ; Search time 12.7987 Seconds  
(without alignments)  
181.822 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_532\_586

Perfect score: 332

Sequence: 1 CSLSNGAHCSGDCPCNNNTSC.....CDITEYCTGDSGOCPPNLHK 55

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 323717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA.\*  
2: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	100.0	696	US-09-351-414-2	Sequence 2, Appli
2	213	64.2	488	US-08-243-542-1	Sequence 1, Appli
3	213	64.2	488	US-08-477-407-1	Sequence 1, Appli
4	213	64.2	488	US-08-484-355-1	Sequence 1, Appli
5	213	64.2	524	US-08-243-542-2	Sequence 2, Appli
6	213	64.2	524	US-08-477-407-2	Sequence 2, Appli
7	213	64.2	524	US-08-484-355-2	Sequence 2, Appli
8	213	64.2	670	US-08-243-542-3	Sequence 3, Appli
9	213	64.2	670	US-08-477-407-3	Sequence 3, Appli
10	213	64.2	670	US-08-484-355-3	Sequence 3, Appli
11	213	64.2	769	US-08-243-542-4	Sequence 4, Appli
12	213	64.2	769	US-08-477-407-4	Sequence 4, Appli
13	213	64.2	769	US-08-484-355-4	Sequence 4, Appli
14	168.5	50.8	814	US-09-813-819-4	Sequence 4, Appli
15	168.5	50.8	814	US-09-920-048-4	Sequence 4, Appli
16	168.5	50.8	855	US-09-813-819-2	Sequence 2, Appli
17	168.5	50.8	855	US-09-920-048-2	Sequence 2, Appli
18	160	48.2	802	US-09-632-098-2	Sequence 2, Appli
19	160	48.2	812	US-09-632-098-4	Sequence 2, Appli
20	152	45.8	651	US-08-264-101-2	Sequence 2, Appli
21	152	45.8	651	US-08-765-243-2	Sequence 2, Appli
22	152	45.8	651	PCT-US95-07295-2	Sequence 2, Appli
23	152	45.8	734	US-08-765-243-8	Sequence 8, Appli
24	152	45.8	734	PCT-US95-07295-8	Sequence 8, Appli
25	147	44.3	616	US-09-606-790-1	Sequence 1, Appli
26	145	43.7	722	US-09-617-145-2	Sequence 2, Appli
27	144.5	43.5	781	US-09-738-946-8	Sequence 8, Appli

28	144	43.4	457	1	US-08-264-101-4	Sequence 4, Appli
29	144	43.4	457	2	US-08-765-243-4	Sequence 4, Appli
30	144	43.4	457	5	PCT-US95-07295-4	Sequence 4, Appli
31	144	43.4	735	2	US-08-765-243-6	Sequence 6, Appli
32	144	43.4	735	5	PCT-US95-07295-6	Sequence 6, Appli
33	144	43.4	735	5	PCT-US95-07295-6	Sequence 6, Appli
34	141	42.5	50	5	PCT-US95-05640-16	Sequence 16, Appli
35	141	42.5	50	5	PCT-US95-05640-25	Sequence 25, Appli
36	136	41.0	621	4	US-09-026-001A-6	Sequence 6, Appli
37	132.5	39.9	681	2	US-08-655-345-4	Sequence 4, Appli
38	132.5	39.9	681	3	US-09-183-275-4	Sequence 4, Appli
39	132.5	39.9	681	5	PCT-US96-08407-4	Sequence 4, Appli
40	132.5	39.9	698	4	US-09-727-169-4	Sequence 4, Appli
41	132.5	39.9	698	4	US-09-579-766A-4	Sequence 4, Appli
42	132.5	39.9	698	4	US-09-726-968A-4	Sequence 4, Appli
43	132.5	39.9	807	2	US-08-655-345-2	Sequence 2, Appli
44	132.5	39.9	807	3	US-09-183-275-2	Sequence 2, Appli
45	132.5	39.9	807	5	PCT-US96-08407-2	Sequence 2, Appli
			824	4	US-09-727-169-2	Sequence 2, Appli

## ALIGNMENTS

```

RESULT 1
US-09-351-414-2
; Sequence 2, Application US/09351414
; Patent No. 6265199
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Balindur, Nand
; APPLICANT: Delsheer, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/351,414
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-351-414-2

Query Match      100.0%; Score 332; DB 3; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CSLSNGAHCSGDCPCNNNTSCLFQPRGYECRDVAWECDDITEYCTGDSGOCPPNLHK 55
Db      409 CSLSNGAHCSGDCPCNNNTSCLFQPRGYECRDVAWECDDITEYCTGDSGOCPPNLHK 463

RESULT 2
US-08-243-542-1
; Sequence 1, Application US/08243542
; Patent No. 5552526
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: : EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P. C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible

```

OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 488 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain CDNA library  
US-08-243-542-1

Query Match 64.2%; Score 213; DB 1; Length 488;  
Best Local Similarity 65.5%; Pred. No. 6.6e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGACSDGPCNNNTSCLFQPRGYECRDVNECDITTEYCTGSGCCPNLHK 55  
Db 371 CTLTHDAMCSGDLCCR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPNLHK 423

RESULT 3  
US-08-477-407-1  
Sequence 1, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542

FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 488 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain CDNA library  
US-08-477-407-1

Query Match 64.2%; Score 213; DB 1; Length 488;  
Best Local Similarity 65.5%; Pred. No. 6.6e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGACSDGPCNNNTSCLFQPRGYECRDVNECDITTEYCTGSGCCPNLHK 55  
Db 371 CTLTHDAMCSGDLCCR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPNLHK 423

RESULT 4  
US-08-484-355-1  
Sequence 1, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 488 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-484-355-1

Query Match 64.2%; Score 213; DB 1; Length 488;  
Best Local Similarity 65.5%; Pred. No. 6,6e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGHCSGPGCCNNTSCLPFGRGYCRDAVNECDITEYCTGSGCCPPLHK 55  
Db 371 CILTHDMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPLHK 423

## RESULT 5

US-08-243-542-2  
Sequence 2, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 524 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-243-542-2

Query Match 64.2%; Score 213; DB 1; Length 524;  
Best Local Similarity 65.5%; Pred. No. 7e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLSNGHCSGPGCCNNTSCLPFGRGYCRDAVNECDITEYCTGSGCCPPLHK 55  
Db 378 CILTHDMCSDGLCCRR--CKYEPBGVSCREAVNECDIAETCTGDSGCCPPLHK 430

## RESULT 6

US-08-477-407-2  
Sequence 2, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 524 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain CDNA library  
US-08-477-407-2

Query Match 64.2%; Score 213; DB 1; Length 524;  
Best Local Similarity 65.5%; Pred. No. 7e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Qy 1 CSLSNAGHSDGPCNNNTSCLEFOPRGYECRDVNECDITEYCTGDSGCCPPLMK 55  
Db 378 CULTHDMCSDGLCCRR--CKYPRGVSCREAVNECDIAETCTGDSGCCPPLMK 430

## RESULT 7

US-08-484-355-2  
Sequence 2, Application US/08484355  
Patent No. 5705341

GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terrence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 524 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens

IMMEDIATE SOURCE:  
LIBRARY: human fetal brain CDNA library  
US-08-484-355-2

Query Match 64.2%; Score 213; DB 1; Length 524;  
Best Local Similarity 65.5%; Pred. No. 7e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Qy 1 CSLSNAGHSDGPCNNNTSCLEFOPRGYECRDVNECDITEYCTGDSGCCPPLMK 55  
Db 378 CULTHDMCSDGLCCRR--CKYPRGVSCREAVNECDIAETCTGDSGCCPPLMK 430

## RESULT 8

US-08-243-542-3  
Sequence 3, Application US/08243542  
Patent No. 5552526

GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terrence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 670 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens

IMMEDIATE SOURCE:  
LIBRARY: human fetal brain CDNA library  
US-08-243-542-3

Query Match 64.2%; Score 213; DB 1; Length 670;  
Best Local Similarity 65.5%; Pred. No. 8.8e-14;  
Matches 36; Conservative 6; Mismatches 11; Indels 2; Gaps 1;



OY 1 CSLNAGACSDGPCNNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55  
Db 378 CTLTHDAMCSDGLCCRR--CKYERGVSCRAVNECDIAETCTGDSGCCPPNLHK 430

## RESULT 9

US-08-477-407-3  
Sequence 3, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAS  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teruyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 670 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-477-407-3

Query Match 64.2%; Score 213; DB 1; Length 670;

Best Local Similarity 65.5%; Pred. No. 8.8e-14; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLNAGACSDGPCNNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55  
Db 378 CTLTHDAMCSDGLCCRR--CKYERGVSCRAVNECDIAETCTGDSGCCPPNLHK 430

RESULT 10

US-08-484-355-3  
Sequence 3, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAS  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teruyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 670 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-484-355-3

Query Match 64.2%; Score 213; DB 1; Length 670;

Best Local Similarity 65.5%; Pred. No. 8.8e-14; Mismatches 11; Indels 2; Gaps 1;

OY 1 CSLNAGACSDGPCNNNTSCLFOPRGYECRDVAVNECDITEYCTGDSGCCPPNLHK 55  
Db 378 CTLTHDAMCSDGLCCRR--CKYERGVSCRAVNECDIAETCTGDSGCCPPNLHK 430

Search completed: October 21, 2003, 10:12:26  
Job time : 13.7987 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 10:06:05 ; Search time 24.2013 Seconds  
(without alignments)  
181.822 Million cell updates/sec

Title: US-09-634-252A-4\_COPY\_496\_599

Perfect score: 630  
Sequence: 1 PTKLFEPTECGNGYVAGEE.....CPNLHKODGYACNONOGR 104

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 323717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	630	100.0	696	3	US-09-351-414-2
2	404.5	64.2	488	1	US-08-243-542-1
3	404.5	64.2	488	1	US-08-477-407-1
4	404.5	64.2	488	1	US-08-484-355-1
5	404.5	64.2	524	1	US-08-243-542-2
6	404.5	64.2	524	1	US-08-477-407-2
7	404.5	64.2	524	1	US-08-484-355-2
8	404.5	64.2	670	1	US-08-243-542-3
9	404.5	64.2	670	1	US-08-477-407-3
10	404.5	64.2	769	1	US-08-484-355-3
11	404.5	64.2	769	1	US-08-243-542-4
12	404.5	64.2	769	1	US-08-477-407-4
13	404.5	64.2	769	1	US-08-484-355-4
14	292	46.3	802	4	US-09-632-098-2
15	292	46.3	812	4	US-09-632-098-4
16	276.5	43.9	781	4	US-09-738-946-8
17	274	43.5	814	3	US-09-813-819-4
18	274	43.5	814	4	US-09-920-048-4
19	274	43.5	855	3	US-09-813-819-2
20	274	43.5	855	4	US-09-920-048-2
21	263	41.7	621	4	US-09-026-001A-6
22	259	41.1	613	4	US-09-026-001A-10
23	259	41.1	621	4	US-09-026-001A-18
24	257	40.8	751	2	US-08-836-443-3
25	253.5	40.2	651	1	US-08-264-101-2
26	253.5	40.2	651	1	US-08-765-243-2
27	253.5	40.2	651	5	PCT-US95-07295-2

28	253.5	40.2	734	2	US-08-765-243-8	Sequence 8, Appli
29	253.5	40.2	734	5	PCT-US95-07295-8	Sequence 8, Appli
30	252	40.0	616	4	US-09-608-790-1	Sequence 1, Appli
31	242	38.4	722	4	US-09-617-145-2	Sequence 2, Appli
32	215.5	34.2	462	4	US-09-026-001A-16	Sequence 16, Appli
33	209.5	33.3	457	1	US-08-264-101-4	Sequence 4, Appli
34	209.5	33.3	457	2	US-08-765-243-4	Sequence 4, Appli
35	209.5	33.3	457	5	PCT-US95-07295-4	Sequence 4, Appli
36	209.5	33.3	735	2	US-08-765-243-6	Sequence 6, Appli
37	209.5	33.3	735	5	PCT-US95-07295-6	Sequence 6, Appli
38	207.5	32.9	521	4	US-09-026-001A-12	Sequence 12, Appli
39	207.5	32.9	592	4	US-09-026-001A-14	Sequence 14, Appli
40	207	32.9	1239	2	US-08-937-931-2	Sequence 2, Appli
41	207	32.9	1239	3	US-09-285-502-2	Sequence 2, Appli
42	207	32.9	1239	4	US-09-709-126-2	Sequence 2, Appli
43	207	32.9	1239	4	US-09-871-385A-2	Sequence 2, Appli
44	200.5	31.8	529	2	US-08-836-442-3	Sequence 3, Appli
45	200.5	31.8	749	2	US-08-937-931-8	Sequence 8, Appli

ALIGNMENTS

```
RESULT 1
US-09-351-414-2
; Sequence 2, Application US/09351414
; Patent No. 6265199
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Balindur, Nand
; APPLICANT: Delisher, Theresa A.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: DISTTEGRIN HOMOLOG
; FILE REFERENCE: 98-29
; CURRENT APPLICATION NUMBER: US/09/351,414
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-351-414-2

Query Match 100.0%; Score 630; DB 3; length 696;
Best Local Similarity 100.0%; Pred. No. 7.7e-45;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTKLFEPTECGNGYVAGEEBCDCGFHYECYGLCKKCSLSNGAHCSDGPCCNNTSCLP 60
Db 373 PTKLFEPTECGNGYVAGEEBCDCGFHYECYGLCKKCSLSNGAHCSDGPCCNNTSCLP 432

Qy 61 RGYECRDVNECDITTEYCTGDSGCGCPNLHKODGYACNONOGR 104
Db 433 RGYECRDVNECDITTEYCTGDSGCGCPNLHKODGYACNONOGR 476

RESULT 2
US-08-243-542-1
; Sequence 1, Application US/08243542
; Patent No. 5552526
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
```

ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
CLASSIFICATION: 530  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teriyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 488 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-243-542-1

Query Match 64.2%; Score 404.5; DB 1; Length 488;  
Best Local Similarity 65.4%; Pred. No. 2.3e-26;  
Matches 70; Conservative 9; Mismatches 23; Indels 5; Gaps 2;

QY 1 PTKLEPTGCGNGVYAGEECDCGFHVECY--GLCKKCSLSNGAHCSGDPCCNNTSCL 57  
DB 332 PLKLDPEPCGNGFVAGEECDCGSGVQCSRAGNCKCKKTLTHDAMCSDGLCCRR--CK 389  
QY 58 FQPRGVECDVAVNECDITFYCTGDSGQCPPLNKHODGYACNONGRC 104  
DB 390 YEPGVSCEAVNECDIAETCTGDSGQCPPLNKHLDGYVCDHEQGR 436

RESULT 3  
US-08-477-407-1  
Sequence 1, Application US/08477407  
Patent No. 561351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible

OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teriyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 488 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-477-407-1

Query Match 64.2%; Score 404.5; DB 1; Length 488;  
Best Local Similarity 65.4%; Pred. No. 2.3e-26;  
Matches 70; Conservative 9; Mismatches 23; Indels 5; Gaps 2;

QY 1 PTKLEPTGCGNGVYAGEECDCGFHVECY--GLCKKCSLSNGAHCSGDPCCNNTSCL 57  
DB 332 PLKLDPEPCGNGFVAGEECDCGSGVQCSRAGNCKCKKTLTHDAMCSDGLCCRR--CK 389  
QY 58 FQPRGVECDVAVNECDITFYCTGDSGQCPPLNKHODGYACNONGRC 104  
DB 390 YEPGVSCEAVNECDIAETCTGDSGQCPPLNKHLDGYVCDHEQGR 436

RESULT 4  
US-08-484-355-1  
Sequence 1, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0



RESULT 8 -  
 US-08-243-542-3  
 : Sequence 3, Application US/08243542  
 : Patent No. 5552526  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: NAKAMURA, YUSUKE  
 :  
 : APPLICANT: EMI, MITSURU  
 :  
 : TITLE OF INVENTION: MDC PROTEINS AND DNAS  
 :  
 : TITLE OF INVENTION: ENCODING THE SAME  
 :  
 : NUMBER OF SEQUENCES: 20  
 :  
 : CORRESPONDENCE ADDRESS:  
 :  
 : ADDRESSEE: FLYNN, THEIL, BOUTELL & TANIS P.C.  
 :  
 : STREET: 2026 Rambling Road  
 :  
 : CITY: Kalamazoo  
 :  
 : STATE: Michigan  
 :  
 : COUNTRY: USA  
 :  
 : ZIP: 49008-1699  
 :  
 : COMPUTER READABLE FORM:  
 :  
 : MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
 :  
 : COMPUTER: IBM PC/XT/AT Compatible  
 :  
 : OPERATING SYSTEM: MS-DOS 5.0  
 :  
 : SOFTWARE: Wordperfect 5.0  
 :  
 : CURRENT APPLICATION DATA:  
 :  
 : APPLICATION NUMBER: US/08/243,542  
 :  
 : FILING DATE:  
 :  
 : CLASSIFICATION: 530  
 :  
 : PRIOR APPLICATION DATA:  
 :  
 : APPLICATION NUMBER: JP 5-136602  
 :  
 : FILING DATE: 14 MAY 1993  
 :  
 : APPLICATION NUMBER: JP 5-257455  
 :  
 : FILING DATE: 22 SEPTEMBER 1993  
 :  
 : APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teriyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 670 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-243-542-3

Query Match 64.2%; Score 404.5; DB 1; Length 670;  
Best Local Similarity 65.4%; Pred. No. 3.1e-26;  
Matches 70; Conservative 9; Mismatches 23; Indels 5; Gaps 2;  
QY 1 PRTLFPTECGNGYVAGEECDCGPFVVECT---GLCKKCSLSNGAHGSDGPCNNNTSCL 57  
DB 339 PLKLDLPBECGNGFVAGEECDCGVSQVBCSRAGNCCCKCTLTHDMCSDGLCCRR--CK 396  
QY 58 FQPRGYECRDVAVNECDITEYCTGDSGQCPNHLKHODGYACNONOGR 104  
DB 397 YEPRGVSCRAVNECDIAETCTGDSQCPNHLKLDGYICDHEQGR 443

RESULT 9  
US-08-477-407-3  
Sequence 3, Application US/08477407  
Patent No. 563351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328

FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teriyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 670 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
US-08-477-407-3

Query Match 64.2%; Score 404.5; DB 1; Length 670;  
Best Local Similarity 65.4%; Pred. No. 3.1e-26;  
Matches 70; Conservative 9; Mismatches 23; Indels 5; Gaps 2;  
QY 1 PRTLFPTECGNGYVAGEECDCGPFVVECT---GLCKKCSLSNGAHGSDGPCNNNTSCL 57  
DB 339 PLKLDLPBECGNGFVAGEECDCGVSQVBCSRAGNCCCKCTLTHDMCSDGLCCRR--CK 396  
QY 58 FQPRGYECRDVAVNECDITEYCTGDSGQCPNHLKHODGYACNONOGR 104  
DB 397 YEPRGVSCRAVNECDIAETCTGDSQCPNHLKLDGYICDHEQGR 443

RESULT 10  
US-08-484-355-3  
Sequence 3, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: Wordperfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Terryence F. Chapman  
 REGISTRATION NUMBER: 32 549  
 REFERENCE/DOCKET NUMBER: Puruya Case 1313  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (616) 381-1156  
 TELEFAX: (616) 381-5465  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 670 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 IMMEDIATE SOURCE:  
 LIBRARY: human fetal brain cDNA library  
 US-08-484-355-3

Query Match 64.2%; Score 404.5; DB 1; Length 670;  
 Best Local Similarity 65.4%; Pred. No. 3,1e-26;  
 Matches 70; Conservative 9; Mismatches 23; Indels 5; Gaps 2;  
 Oy 1 PTKLPETECNGYVEAGEECDCGFHVECY--GLCKKCSLSNGAHGSDGPPCCNNTSCL 57  
 Db 339 PLKLLDPECCNGGFVEAGEECDCGSGVQECSPRAGNCKKCTLTTHDAMCSDGLCCRR--CK 396  
 Oy 58 FQPRGYEGRDAVNECDITEYCTGDSGCGCPNLHKQDGYACNQNQRC 104  
 Db 397 YEPKGVSCREAVNECDIAETCTGDSGCGCPNLHKLDGYCDHEQGR 443

Search completed: October 21, 2003, 10:12:25  
 Job time : 25.2013 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 21, 2003, 10:20:51 ; Search time 62.1384 Seconds  
(without alignments)  
738.736 Million cell updates/sec

Title: US-09-634-252a-4\_COPY\_496\_599  
Perfect score: 630  
Sequence: 1 PTLFEPTEGNGYVEAGE.....CPPLHKQDQVACNQNGRC 104

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues  
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh  
-Q=/cgn2\_1/USFTO.spool/US0963452/rnat\_21102003\_090325\_34C1/app\_query.fasta\_1.462  
-DB=Issued\_Patents\_NA -OPMT=faetap -SUFFIX=p2n.rni -MIMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=10  
-MODE=LOCAL -OUTPMT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USERS=US0963452.@CNC 1.1.76@rnat\_21102003\_090325\_3401 -NCPU=6 -ICPU=3  
-NO\_MAP -LARSBOUNDARY -NEG\_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued Patents NA.\*  
1: /cgn2\_6/p2data/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/p2data/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/p2data/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/p2data/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/p2data/2/ina/6C.COMB.seq.\*  
6: /cgn2\_6/p2data/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	2268	3	US-09-351-414-1 Sequence 1, Appli
2	567	90.0	2088	3	US-09-351-414-3 Sequence 3, Appli
3	404.5	64.2	1464	1	US-08-243-542-5 Sequence 5, Appli
4	404.5	64.2	1464	1	US-08-477-407-5 Sequence 5, Appli
5	404.5	64.2	1464	1	US-08-484-355-5 Sequence 5, Appli
6	404.5	64.2	2913	1	US-08-243-542-7 Sequence 7, Appli
7	404.5	64.2	2913	1	US-08-477-407-7 Sequence 7, Appli
8	404.5	64.2	2913	1	US-08-484-355-7 Sequence 7, Appli
9	404.5	64.2	2923	1	US-08-243-542-6 Sequence 6, Appli
10	404.5	64.2	2923	1	US-08-477-407-6 Sequence 6, Appli
11	404.5	64.2	2923	1	US-08-484-355-6 Sequence 6, Appli
12	404.5	64.2	3183	1	US-08-243-542-8 Sequence 8, Appli

13	404.5	64.2	3183	1	US-08-477-407-8 Sequence 8, Appli
14	404.5	64.2	3183	1	US-08-484-355-8 Sequence 8, Appli
15	292	46.3	3431	1	US-09-632-098-1 Sequence 1, Appli
16	292	46.3	3468	4	US-09-632-098-3 Sequence 3, Appli
17	276.5	43.9	3984	4	US-09-738-946-7 Sequence 7, Appli
18	274	43.5	2968	3	US-09-813-819-1 Sequence 1, Appli
19	274	43.5	2968	4	US-09-920-048-1 Sequence 1, Appli
20	265	42.1	2406	4	US-09-632-098-5 Sequence 5, Appli
21	265	42.1	2439	4	US-09-632-098-6 Sequence 6, Appli
22	263	41.7	2050	4	US-09-632-098-5 Sequence 5, Appli
23	259	41.1	2335	4	US-09-632-098-5 Sequence 5, Appli
24	259	41.1	2339	4	US-09-632-098-5 Sequence 5, Appli
25	257	40.8	2251	2	US-08-836-443-2 Sequence 2, Appli
26	257	40.8	2648	2	US-08-836-443-1 Sequence 1, Appli
27	255	40.5	1463	5	PCT-US93-05640-10 Sequence 10, Appli
28	253.5	40.2	2373	1	US-08-765-243-1 Sequence 1, Appli
29	253.5	40.2	2373	2	US-08-765-243-1 Sequence 1, Appli
30	253.5	40.2	2373	5	PCT-US95-07295-1 Sequence 1, Appli
31	253.5	40.2	2650	2	US-08-765-243-7 Sequence 7, Appli
32	253.5	40.2	2650	5	PCT-US95-07295-7 Sequence 7, Appli
33	252	40.0	1851	4	US-09-608-790-2 Sequence 2, Appli
34	242	38.4	2169	4	US-09-617-145-1 Sequence 1, Appli
35	215.5	34.2	1820	4	US-09-626-001A-15 Sequence 15, Appli
36	211	33.5	1550	2	US-08-836-442-2 Sequence 2, Appli
37	211	33.5	2056	2	US-08-836-442-1 Sequence 1, Appli
38	209.5	33.3	1768	1	US-08-765-243-3 Sequence 3, Appli
39	209.5	33.3	1768	2	US-08-765-243-3 Sequence 3, Appli
40	209.5	33.3	1768	5	PCT-US95-07295-3 Sequence 3, Appli
41	209.5	33.3	2553	2	US-08-765-243-5 Sequence 5, Appli
42	209.5	33.3	2553	5	PCT-US95-07295-5 Sequence 5, Appli
43	207.5	32.9	2288	4	US-09-626-001A-11 Sequence 11, Appli
44	207.5	32.9	2309	4	US-09-626-001A-13 Sequence 13, Appli
45	207	32.9	5630	2	US-08-937-931-1 Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-351-414-1  
Sequence 1, Application US/09351414  
Patent No. 6265199  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Beinhard, Nand  
APPLICANT: Deisher, Theresa A.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
FILE REFERENCE: 98-29  
CURRENT APPLICATION NUMBER: US/09/351,414  
CURRENT FILING DATE: 1999-07-09  
NUMBER OF SEQ. ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)...(2090)  
NAME/KEY: misc\_feature  
LOCATION: (1)...(2268)  
OTHER INFORMATION: n = A,T,C or G  
US-09-351-414-1

## Alignment Scores:

Pred. No.: 2,17e-63 Length: 2268  
Score: 630.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-351-414-1 (1-2268)

QY 1 ProThrIysLeuPheGluProThrGluCysGlyAsnGlyTYrValGluAlaGlyGluGlu 20  
DB 1119 CCAACAAAGCTATTGAGCCCAACGGAATGTGAATGATAGCTGGAAGCTGGAGAGAG 1178  
QY 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysValGlyLysCysSerLeuSer 40  
DB 1179 TGTGATTTGTGTTTCATGATGTAATGCTATGATATGCTGTAAGAAATGTTCCCTCTCC 1238  
QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGluPro 60  
DB 1229 AACGGGGCTCACTGACGACGACCGGCGCTGCTGTAACATACCTCATGCTTTTTCAGCCA 1298  
QY 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGluTYrCysThrGly 80  
DB 1299 CGAGGGTATGAATGCCGGGATGCTGGAACGAGTGTGATATTACTGAATATTGTACTGGA 1358  
QY 81 AspSerGlyGlnCysProProAsnLeuHisIlysgInAspGlyTYrAlaCysAsnGlnAsn 100  
DB 1359 GACTCTGTGACAGTGCACCAAAATCTTCATTAAGCAAGCGATATGCAATGCAAAAT 1418  
QY 101 GlnGlyArgCys 104  
DB 1419 CAGGGCCGCTGC 1430  
QY 1419 CAGGGCCGCTGC 1430  
RESULT 2  
US-09-351-414-3  
; Sequence 3, Application US/09351414  
; Patent No. 6265199  
; GENERAL INFORMATION:  
; APPLICANT: Shepard, Paul O.  
; APPLICANT: Baidour, Nand  
; APPLICANT: Deisher, Theresa A.  
; APPLICANT: Bishop, Paul D.  
; TITLE OF INVENTION: DISINTEGRIN HOMOLOG  
; FILE REFERENCE: 98-29  
; CURRENT APPLICATION NUMBER: US/09/351,414  
; CURRENT FILING DATE: 1999-07-09  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 3  
; LENGTH: 2088  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: zdn1 amino acid degenerate sequence  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (1)...(2088)  
; OTHER INFORMATION: n is any nucleotide  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)\_(2088)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-351-414-3  
Alignment Scores:  
Pred. No.: 3,77e-56 Length: 2088  
Score: 567.00 Matches: 91  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 13  
Query Match: 90.00% Indels: 0  
DB: 3 Gaps: 0  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-09-351-414-3 (1-2088)  
QY 1 ProThrIysLeuPheGluProThrGluCysGlyAsnGlyTYrValGluAlaGlyGluGlu 20  
DB 1117 CCNACNAARYTNTTTCGARCNCNCGARCTGCGNAAYGNTATGTAAGCNCNGNARGAR 1176  
QY 21 CysAspCysGlyPheHisValGluCysTYrGlyLeuCysValGlyLysCysSerLeuSer 40  
DB 1177 TGTGATTTGTGTTTCATGATGTAATGCTATGATATGCTGTAAGAAATGTTCCCTCTCC 1236

QY 41 AsnGlyAlaHisCysSerAspGlyProCysCysAsnAsnThrSerCysLeuPheGluPro 60  
DB 1227 AATGNGNCNCAYTGTGWSGATGCGNCNTGTGYAAAYAAACWMSNTGTYTNTTTCARCCN 1296  
QY 61 ArgGlyTYrGluCysArgAspAlaValAsnGluCysAspIleThrGluTYrCysThrGly 80  
DB 1297 MGNGNTAYGARTGTMGNGAYGCGNTNAAYGARTGYGAYATHACNGARTAYTGVAACNGN 1356  
QY 81 AspSerGlyGlnCysProProAsnLeuHisIlysgInAspGlyTYrAlaCysAsnGlnAsn 100  
DB 1357 GAYWSNGNCARTGTCNCNCAAYTNCAYARCARGAYGNTAYGCVTGYAAACARAA 1416  
QY 101 GlnGlyArgCys 104  
DB 1417 CARGSNMNTGY 1428  
RESULT 3  
US-08-243-542-5  
; Sequence 5, Application US/08243542  
; Patent No. 5552526  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: EMI, MITSURU  
; TITLE OF INVENTION: MDC PROTEINS AND DNAs  
; TITLE OF INVENTION: ENCODING THE SAME  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLYNN, THEL, BOUTELL & TANIS P.C.  
; STREET: 2026 Rambling Road  
; CITY: Kalamazoo  
; STATE: Michigan  
; COUNTRY: USA  
; ZIP: 49008-1699  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT Compatible  
; OPERATING SYSTEM: MS-DOS 5.0  
; SOFTWARE: WordPerfect 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/243,542  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-136602  
; FILING DATE: 14 MAY 1993  
; APPLICATION NUMBER: JP 5-257455  
; FILING DATE: 22 SEPTEMBER 1993  
; APPLICATION NUMBER: JP 6-49904  
; FILING DATE: 23 FEBRUARY 1994  
; APPLICATION NUMBER: JP 6-73328  
; FILING DATE: 12 APRIL 1994  
; APPLICATION NUMBER: JP 6-84470  
; FILING DATE: 22 APRIL 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Terryence F. Chapman  
; REGISTRATION NUMBER: 32 549  
; REFERENCE/DOCKET NUMBER: Furuya Case 1313  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (616) 381-1156  
; TELEFAX: (616) 381-5465  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1464 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; IMMEDIATE SOURCE:  
; LIBRARY: human fetal brain cdna library  
; FEATURE:

NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-243-542-5

Alignment Scores:  
Pred. No.: 1,47e-37 Length: 1464  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-08-243-542-5 (1-1464)

Qy 1 ProthrlrlysluPheglupProthrgluCyaglyYsnglyTyYrValGluAlaGlyGluGlu 20  
Db 994 CCCCTCAAGCTCTCTGAGACCCCGACAGTGGGGAACGGCTTCGTGAGAGCGGGGAGAG 1053

Qy 21 CysAspCysgIyPhehlsvalGluCysTyr-----GlyLeuCyCySlySlyCys 37  
Db 1054 TGGACATGGCGGCTGGTGACAGAGTGGACGCCGCGAGGTGGCAATGCTGCAAGAAATGC 1113

Qy 38 SerleuSeranGlyAlahlsCysSerAspGlyProCyCySasnaenThrSerCysleu 57  
Db 1114 ACCCTGACTCAGACGCGCATGTGCAGCGACGCGGCTGTGCTGCG:CGC-----TGCAG 1167

Qy 58 PheGlnProArGlyTyYrGluCyArGAspAlaValaenGluCyAspIleThrGluTyr 77  
Db 1168 TACCAACACCGGGGTGTCTCTGCGAGAGCGCGGAAACGAGTGCACATGCGGAGACC 1227

Qy 78 CysThrGlyAspSerGlyGlnCysProProAsnleuHlslySglnaSfGlyTyYrAlaCys 97  
Db 1228 TGCACCGGGGACTCTAGCCAGTGGCCGCTTAACCTGCACAGCTGACAGGTTACTACTGT 1287

Qy 98 AsnGlnaenGlnGlyArGys 104  
Db 1288 GACCATGAGCAGGCGCGCTGC 1308

RESULT 4  
US-08-477-407-5  
Sequence 5, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477.407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243.542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Teriyence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-477-407-5

Alignment Scores:  
Pred. No.: 1,47e-37 Length: 1464  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-08-477-407-5 (1-1464)

Qy 1 ProthrlrlysluPheglupProthrgluCyaglyYsnglyTyYrValGluAlaGlyGluGlu 20  
Db 994 CCCCTCAAGCTCTCTGAGACCCCGACAGTGGGGAACGGCTTCGTGAGAGCGGGGAGAG 1053

Qy 21 CysAspCysgIyPhehlsvalGluCysTyr-----GlyLeuCyCySlySlyCys 37  
Db 1054 TGGACATGGCGGCTGGTGACAGAGTGGACGCCGCGAGGTGGCAATGCTGCAAGAAATGC 1113

Qy 38 SerleuSeranGlyAlahlsCysSerAspGlyProCyCySasnaenThrSerCysleu 57  
Db 1114 ACCCTGACTCAGACGCGCATGTGCAGCGACGCGGCTGTGCTGCGCGC-----TGCAG 1167

Qy 58 PheGlnProArGlyTyYrGluCyArGAspAlaValaenGluCyAspIleThrGluTyr 77  
Db 1168 TACCAACACCGGGGTGTCTCTGCGAGAGCGCGTGAACAGTGCACATGCGGAGACC 1227

Qy 78 CysThrGlyAspSerGlyGlnCysProProAsnleuHlslySglnaSfGlyTyYrAlaCys 97  
Db 1228 TGCACCGGGGACTCTAGCCAGTGGCCGCTTAACCTGCACAGCTGACAGGTTACTACTGT 1287

Qy 98 AsnGlnaenGlnGlyArGys 104  
Db 1288 GACCATGAGCAGGCGCGCTGC 1308

RESULT 5  
US-08-484-355-5  
Sequence 5, Application US/08484355  
Patent No. 5705341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.

STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Futuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1464  
US-08-484-355-5  
Alignment Scores:  
Pred. No.: 1.47e-37 Length: 1464  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2  
US-09-634-252a-4\_copy\_496\_599 (1-104) x US-08-484-355-5 (1-1464)  
QY 1 P r o t h r i y l s l e u P h e g l u P r o t h r g l u C y s g l y a s n g l y t y r v a l g l u a l a g l u g l u 20  
DB 994 C C C C T C A A G C T C T C T G A G A G C C C C C A G A G T C G G G A A C G G C T T C T G T G A G G C A G G A G A G 1053  
QY 21 C y s A s p C y s g l P h e H i s v a l g l u C y s t y r ----- G l y l e u C y s C y s l y s l y s C y s 37  
DB 1054 T G C G A C T G C G G C T C G C T G C A G A G T G C A G C C G C G A G G T G C A C T G C T G C A A G A A T G C 1113  
QY 38 S e r l e u S e r A s n g l y A l a H i s C y s S e r A s p g l y P r o C y s C y s A s n A s n t h r S e r C y s l e u 57  
DB 1114 A C C G T A C T C A C A G A G C C A T G T G C A G C A G C G G C T C T G T G T G C G C G C ----- T G C A A G 1167  
QY 58 P h e g l i n P r o A r g g l y t y r t g l u C y s a r g A s p a l a v a l a s n g l u C y s A s p l e t h r i g l u t y r 77

DB 1168 T A C G A A C C A C G G G G T G T C T C T G C C G A G A G C C C G T G A A C G A G T G C G A C A T C G C G A G A C C 1227  
QY 78 C y s t h r g l y A s p s e r g l y G l n C y s p r o P r o b e n l e u H i s l y s g l n A s p g l y t y r A l a c y s 97  
DB 1228 T G C A C C G G G A G A C T C T A G C C A G T G C C C G C T A A C C T G C A C A A G T G C A G C G T T A C T A C T G T 1287  
QY 98 A s n g l n A s n g l n g l y A r g C y s 104  
DB 1288 G A C C A T G A C G A C G G C G C T G C 1308  
RESULT 6  
US-08-243-542-7  
Sequence 7, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: : EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAS  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Futuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2913 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: 5' UTR  
LOCATION: 1..27  
FEATURE:  
NAME/KEY: 3' UTR  
LOCATION: 2038..2913  
FEATURE:

NAME/KEY: CDS  
LOCATION: 28..2037  
US-08-243-542-7

Alignment Scores:  
Pred. No.: 3,74e-37 Length: 2913  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-08-243-542-7 (1-2913)

QY 1 ProThrlYsleuPhheGluProThrgluCyGlyAsnGlyTyrValGluAlaGluGlu 20  
Db 1042 CCCCTCAAGCTCTTGAGACCCCGCAGAGTGGGGAACGGCTTCTGAGAGGAGGAGAG 1101

QY 21 CysAspCysGlyPheHisValGluCysTyr-----GlyLeuCysCysLysLysCys 37  
Db 1102 TGGGACTGCGGCTCGGTGACAGAGTGCAGCGCGCAGAGTGCACACTGCTCAAGAAATGC 1161

QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnThrSerCysLeu 57  
Db 1162 ACCCTGACTCAGCAGCGCATGTGCAGCAGCGGCTGTGTCGCCG-----TGCAG 1215

QY 58 PheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrgluTyr 77  
Db 1216 TACGAACACAGGGGTGTGCTCTCCGAGAGCGCTGACACAGTGCACATCGCGAGACC 1275

QY 78 CysThrGlyAspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTyrAlaCys 97  
Db 1276 TGCACCGGGGACTCTAGCCAGTGCAGCGCGCTAACCTGCACAAAGTGCAGCGTTACTACTGT 1335

QY 98 AsnGlnAsnGlnGlyArgCys 104  
Db 1336 GACCATGAGCAGGCGCGCTGC 1356

RESULT 7  
US-08-477-407-7  
Sequence 7, Application US/08477407  
Patent No. 5631351  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: : EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAs  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
City: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,407  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terrence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Furuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2913 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library

FEATURE:  
NAME/KEY: 5' UTR  
LOCATION: 1..27  
FEATURE:  
NAME/KEY: 3' UTR  
LOCATION: 2038..2913  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..2037  
US-08-477-407-7

Alignment Scores:  
Pred. No.: 3,74e-37 Length: 2913  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2

US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-08-477-407-7 (1-2913)

QY 1 ProThrlYsleuPhheGluProThrgluCyGlyAsnGlyTyrValGluAlaGluGlu 20  
Db 1042 CCCCTCAAGCTCTTGAGACCCCGCAGAGTGGGGAACGGCTTCTGAGAGGAGGAGAG 1101

QY 21 CysAspCysGlyPheHisValGluCysTyr-----GlyLeuCysCysLysLysCys 37  
Db 1102 TGGGACTGCGGCTCGGTGACAGAGTGCAGCGCGCAGAGTGCACACTGCTCAAGAAATGC 1161

QY 38 SerLeuSerAsnGlyAlaHisCysSerAspGlyProCysCysAsnThrSerCysLeu 57  
Db 1162 ACCCTGACTCAGCAGCGCATGTGCAGCAGCGGCTGTGTCGCCG-----TGCAG 1215

QY 58 PheGlnProArgGlyTyrGluCysArgAspAlaValAsnGluCysAspIleThrgluTyr 77  
Db 1216 TACGAACACAGGGGTGTGCTCTCCGAGAGCGCTGACACAGTGCACATCGCGAGACC 1275

QY 78 CysThrGlyAspSerGlyGlnCysProProAsnLeuHisLysGlnAspGlyTyrAlaCys 97  
Db 1276 TGCACCGGGGACTCTAGCCAGTGCAGCGCGCTAACCTGCACAAAGTGCAGCGTTACTACTGT 1335

QY 98 AsnGlnAsnGlnGlyArgCys 104  
Db 1336 GACCATGAGCAGGCGCGCTGC 1356

RESULT 8  
US-08-484-355-7  
Sequence 7, Application US/08484355  
Patent No. 570341  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAS  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THEIL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,355  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/243,542  
FILING DATE: 13-MAY-1994  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Futuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2913 base pairs  
TYPE: nucleic acid  
TOPOLOGY: linear  
STRANDEDNESS: double  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
LIBRARY: human fetal brain cDNA library  
FEATURE:  
NAME/KEY: 5' UTR  
LOCATION: 1..27  
NAME/KEY: 3' UTR  
LOCATION: 2038..2913  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..2037  
US-08-484-355-7  
Alignment Scores:  
Pred. No.: 3.74e-37 Length: 2913  
Score: 404.50 Matches: 70  
Percent Similarity: 73.83% Conservative: 9  
Best Local Similarity: 65.42% Mismatches: 23  
Query Match: 64.21% Indels: 5  
DB: 1 Gaps: 2  
US-09-634-252a-4\_COPY\_496\_599 (1-104) x US-08-484-355-7 (1-2913)  
QY 1 ProthrllysLeupheglupProthrlglucysgllyaenglyTyrrvalglualaglygluclu 20

Db 1042 CCCCTCAAGCTCTCGAGACCCCGGAGAGTCCGAGAGGAGGAGAG 1101  
QY 21 CysAspCysgslgylPheHlsvalgluCystr-----GlyLeuCyCyvalgylsCy 37  
Db 1102 TGCAGCTGGGCTGGGTGCGAGAGTGCAGCCGCCAGGTGCACACTGCTGCAAGAAATGC 1161  
QY 38 SerLeuSerAenglAlahleCySerAaspGlyProCyCyAsnAsnThSerCysleu 57  
Db 1162 ACCCTGACTCAGACCGCATGTGCAGGACCGGCTGCTGCCCGC-----TGCAG 1215  
QY 58 PheglProAtrgglTyrrgluCyAtrgAspAlavalengluCyAspIleThgluTy 77  
Db 1216 TACGACCGGAGCTCTAGCCAGTGCCTTAACTGCAAGCTGAGCGTTACTACTGT 1335  
QY 78 CysThrglAaspsrglgluCyProProhlsnlehlslvsglAaspglyTyrrAlaCy 97  
Db 1276 TGCACCGGGAGCTCTAGCCAGTGCCTTAACTGCAAGCTGAGCGTTACTACTGT 1335  
QY 98 AsnglnAenglnglyAtrgCy 104  
Db 1336 GACCATGACGAGGCGCTGC 1356  
RESULT 9  
US-08-243-542-6  
Sequence 6, Application US/08243542  
Patent No. 5552526  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: EMI, MITSURU  
TITLE OF INVENTION: MDC PROTEINS AND DNAS  
TITLE OF INVENTION: ENCODING THE SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLYNN, THEIL, BOUTELL & TANIS P.C.  
STREET: 2026 Rambling Road  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49008-1699  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage  
COMPUTER: IBM PC/XT/AT Compatible  
OPERATING SYSTEM: MS-DOS 5.0  
SOFTWARE: WordPerfect 5.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/243,542  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-136602  
FILING DATE: 14 MAY 1993  
APPLICATION NUMBER: JP 5-257455  
FILING DATE: 22 SEPTEMBER 1993  
APPLICATION NUMBER: JP 6-49904  
FILING DATE: 23 FEBRUARY 1994  
APPLICATION NUMBER: JP 6-73328  
FILING DATE: 12 APRIL 1994  
APPLICATION NUMBER: JP 6-84470  
FILING DATE: 22 APRIL 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Terryence F. Chapman  
REGISTRATION NUMBER: 32 549  
REFERENCE/DOCKET NUMBER: Futuya Case 1313  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (616) 381-1156  
TELEFAX: (616) 381-5465  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2923 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

```
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
FEATURE:
NAME/KEY: 5' UTR
LOCATION: 1..27
NAME/KEY: 3' UTR
LOCATION: 1600..2923
FEATURE:
NAME/KEY: CDS
LOCATION: 28..1599
US-08-243-542-6

Alignment Scores:
Pred. No.: 3,75e-37 Length: 2923
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 1 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-08-243-542-6 (1..2923)
Qy 1 ProthrylsleuphegipProthrgluCyseglYasnglyTYrValgluAlaglygluGlu 20
Db 1042 CCCCTCAAGCTCTCGAGACCCCGGAGAGTGGCGGAACGGCTTCGTAAGCGCGGAGAG 1101
Qy 21 CysAspCysglYpnehisValgluCyseTYr-----glYleuCyseYsLysCys 37
Db 1102 TGGCACTGGCGGCTCGTGAGAGAGTGCAGCCGCCAGAGTGCACTGCTGCAAGAAATGC 1161
Qy 38 SerleuserAenglyAlahisCyseSerAaspGlyProCyseYsaAuaenThSerCyseleu 57
Db 1162 ACCCTGACTCAGCAGCGCATGTGCAGCGACGGGCTGTGCTGCTCCGCC-----TGCAG 1215
Qy 58 PheglProArnglyTYrGluCyseArGAspAlaValaengluCyseAspIleThrgluTYr 77
Db 1216 TACCAACACAGGGGTGTCTCTCGGAGAGGCGCTGAACGAGTGCAGACATGCGGAGACC 1275
Qy 78 CysThrglyAspSerGlyInCyseProProAenluhIslysglnAspGlyTYrAlaCys 97
Db 1276 TGCACCGGGGACTCTAGCAGGTGCGCGCTAACCCTGCAACAGCTGAGCGTTACTACTGT 1335
Qy 98 AsnglnAsnglnglyA-gCys 104
Db 1336 GACCATGAGCAGGCGCGCTGC 1356

RESULT 10
US-08-477-407-6
Sequence 6, Application US/08477407
Patent No. 5631351
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PLYNN, THIEL, BOUTELL & TANIS P. C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terrylene F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2923 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
FEATURE:
NAME/KEY: 5' UTR
LOCATION: 1..27
FEATURE:
NAME/KEY: 3' UTR
LOCATION: 1600..2923
FEATURE:
NAME/KEY: CDS
LOCATION: 28..1599
US-08-477-407-6

Alignment Scores:
Pred. No.: 3,75e-37 Length: 2923
Score: 404.50 Matches: 70
Percent Similarity: 73.83% Conservative: 9
Best Local Similarity: 65.42% Mismatches: 23
Query Match: 64.21% Indels: 5
DB: 1 Gaps: 2

US-09-634-252a-4_COPY_496_599 (1-104) x US-08-477-407-6 (1..2923)
Qy 1 ProthrylsleuphegipProthrgluCyseglYasnglyTYrValgluAlaglygluGlu 20
Db 1042 CCCCTCAAGCTCTCGAGACCCCGGAGAGTGGCGGAACGGCTTCGTAAGCGCGGAGAG 1101
Qy 21 CysAspCysglYpnehisValgluCyseTYr-----glYleuCyseYsLysCys 37
Db 1102 TGGCACTGGCGGCTCGTGAGAGAGTGCAGCCGCCAGAGTGCACTGCTGCAAGAAATGC 1161
Qy 38 SerleuserAenglyAlahisCyseSerAaspGlyProCyseYsaAuaenThSerCyseleu 57
Db 1162 ACCCTGACTCAGCAGCGCATGTGCAGCGACGGGCTGTGCTGCTCCGCC-----TGCAG 1215
Qy 58 PheglProArnglyTYrGluCyseArGAspAlaValaengluCyseAspIleThrgluTYr 77
Db 1216 TACCAACACAGGGGTGTCTCTCGGAGAGGCGCTGAACGAGTGCAGACATGCGGAGACC 1275
Qy 78 CysThrglyAspSerGlyInCyseProProAenluhIslysglnAspGlyTYrAlaCys 97
```

Db 1276 TGCACCGGGACTTAGCCAGTGCCCCGCTTAAGTGCACAGCTGACGGTTACTACTGT 1335  
Qy 98 AsnglnAsnglnGlyArgCys 104  
::: |||||  
Db 1336 GACCATGAGCAAGGCGCGCTGC 1356

Search completed: October 21, 2003, 12:03:42  
Job time : 68.1384 secs